Research on the Effects of Walking on the Localization Performance of a Human Inspired Perception System for a Biped Humanoid Robot

人間を模擬した知覚系を持つ２足ヒューマノイドロボットの自己位置推定機能に対する歩行動作の影響に関する研究

July 2019

Yukitoshi MINAMI SHIGUEMATSU

南　重松　行紀

Research on the Effects of Walking on the Localization Performance of a Human Inspired Perception System for a Biped Humanoid Robot

人間を模擬した知覚系を持つ２足ヒューマノイドロボットの自己位置推定機能に対する歩行動作の影響に関する研究

July 2019

Waseda University
Graduate School of Advanced Science and Engineering
Department of Integrative Bioscience and Biomedical Engineering, Research on Biorobotics

Yukitoshi MINAMI SHIGUEMATSU
南　重松　行紀

PhD Thesis

Research on the Effects of Walking on the Localization Performance of a Human Inspired Perception System for a Biped Humanoid Robot

Supervisor: Prof. Atsuo Takanishi

Waseda University

Graduate School of Advanced Science and Engineering

Department of Integrative Bioscience and Biomedical Engineering

Takanishi Laboratory, Biped Group

Biorobotics

Minami Shiguematsu, Yukitoshi

5314FE52-9

2019

"One's destination is never a place, but rather a new way of looking at things."

Henry Miller

WASEDA UNIVERSITY

Abstract

Graduate School of Advanced Science and Engineering
Department of Integrative Bioscience and Biomedical Engineering

Doctor of Engineering

Research on the Effects of Walking on the Localization Performance of a Human Inspired Perception System for a Biped Humanoid Robot

by Yukitoshi Minami Shiguematsu

Since we live in a physical world, motion is a requirement to interact with it, as well as with other objects and entities in it. Therefore, being able to track our motions becomes an important ability in order to have satisfactory interactions with the outside world. Localization is the ability to track the position of our bodies while travelling from one place to another, and humans perform it using mainly three sensorial inputs: visual, gravitational/inertial and proprioceptive. The visual input comes from the interpretation of the light absorbed in the photoreceptors inside our eyes and gives us information about the external world. The proprioceptive input gives us information about the relative position and movement of parts of our own body, as well as the strength and effort used for moving, and comes from mechanoreceptors and proprioceptors throughout our body, as well as from muscle spindles and the skin. As for the gravitational and inertial inputs, they mainly come from the vestibular system located in the inner ear of humans, but there are also studies pointing to the fact that humans also have so-called somatic graviceptors, i.e., gravity sensors on their trunks, where the somato-graviceptive sources are thought to be the kidneys and the blood inside large vessels on the body. After the information from the aforementioned senses is obtained, it is combined and fused, changing the weight placed on each sensory input depending on the situation, as there are studies showing, for example, that the visual system performs better at lower frequencies than the vestibular system, but both are integrated in an optimal manner. Clemens et al. (2011) proposed a model called “Indirect pathway model”, in which the information is combined in order to obtain two different estimations: one of the position and orientation of the head, and one of the position and orientation of the body. These estimations are then fused using the kinematic information of the neck. After having obtained the position and orientation estimates of the head and the body, it is then possible

for humans to localize in space while moving through it. There is evidence pointing out that modifying the walking speed has effects on our path integration abilities, making us underestimate distances when walking at slower speeds, as well as walking cadence affecting the performance of path integration, achieving the best performance at about 2 Hz. Also, the human odometer is sensitive to asymmetries in walking style. Motivated by the above facts, we decided to explore the effects of walking gait parameters on the localization of a biped humanoid robot. For this, we first developed a humanlike perception system inspired by the "Indirect path model", which comprised the following subsystems:

1. Head Localization System, using a camera and an inertial measurement unit (IMU) acting as visual and vestibular inputs
2. Body Localization System, using the motor encoders in all the joints and force-torque sensors on the feet, as well as another IMU on the trunk, acting as proprioceptive and somato-graviceptive inputs, respectively.

A loosely coupled Extended Kalman Filter (EKF) based sensor fusion algorithm was used, as it is theorized that is also how humans combine the information from different sensorial sources. The head localization system was first tested in simulation, to verify its effectiveness. Then a first prototype was built, and finally the system was mounted on the head of the humanoid robot WABIAN-2R. For the visual odometry, we focused on sparse odometry algorithms as they are more suited for localization rather than mapping, which suited our purpose. Three different sparse visual odometry (VO) algorithms were tested: a direct, a semi-direct and an indirect algorithm. For the body localization system, preliminary tests were made on the biped humanoid robot WALKMAN, from the Italian Institute of Technology (IIT), to also verify its effectiveness, and subsequently the algorithm was used with the humanoid robot WABIAN-2R. We tested the effects of different human walking parameters on these systems, based on how human localization also changes with these: a. Step length as a proxy for speed and because of its ready applicability to current footstep planners b. Walking style and symmetry For the effects of step length, the direct VO algorithm's performance decreased the longer the step lengths, which along with the analysis of inertial and force/torque data, point to a decrease in performance due to an increase of mechanical vibrations. The indirect VO algorithm's performance decreased in an opposite way, i.e., showing more errors with shorter step lengths, which we show to be due to the effects of drift over time. Finally, the semi-direct VO algorithm showed a performance in-between the previous two. Regarding the walking style and symmetry, changing the walking style from normal to gallop slightly improved the performance of the visual localization, which was related to

a reduction in torques on the feet. Changing the gait temporal symmetry worsened the performance of the visual algorithms, which according to an analysis of inertial data, is related to an increase of mechanical vibrations and camera rotations. Both changes of gait style and symmetry decreased the performance of the kinematic localization, caused by the increase of vertical ground reaction forces, to which kinematic odometry is very sensitive. The observations from both experiments support our claim that gait and footstep planning could be used to improve the performance of localization algorithms in the future. This thesis consists of six chapters in which I present the research background, the proposed perception system with the chosen sensors and algorithms for localization, both for the head and for the body, the experimental setup and results to confirm the effects of walking parameters on the robot's localization, and finally a discussion on limits and possible extensions of this work. The thesis is laid out as follows:

Chapter 1 introduces the research background. More specifically, it presents how humans perceive their motion, briefly describing the different systems participating in it, particularly focusing on the Indirect path model for self-motion estimation, which proposes that humans estimate the motion of their head and their body separately, and then combine this information through the kinematics of the neck. Moreover, it contains the objective of this research and the comparison with other related researches in this field.

Chapter 2 introduces the sensor system proposed for the head motion estimation of the robot, as well as the algorithm to be used to fuse the different sensor inputs. The estimation system comprises visual and gravito-inertial inputs, obtained from a camera and an IMU on the robot's head respectively. The results from the simulation tests, as well as those from the first prototype are presented.

Chapter 3 presents how the walking motions affects the localization performance of the head localization system proposed in the previous chapter. Particularly, step length is the modified walking parameter. Three VO algorithms are compared: a direct, a semi-direct and an indirect VO algorithm. The performance is measured through the absolute trajectory error, and the relative pose error.

Chapter 4 introduces the sensor system proposed for the body motion estimation of the robot, as well as the algorithm to be used to fuse the different sensor inputs for the estimation. The estimation system comprises proprioceptive and gravito-inertial inputs, obtained from joint encoders and force-torque sensors on the feet, as well an IMU on the robot's trunk, respectively. Results of the preliminary experiments with the biped humanoid robot WALKMAN from the IIT are presented.

Chapter 5 presents how the walking motions affects the localization performance of the body localization system proposed in the previous chapter. The modified parameters are step length, walking style and walking symmetry. The performance is measured through the absolute trajectory error, and the relative pose error.

Chapter 6 discusses the quantitative and qualitative results of this work, analyzing the how the localization performance of the proposed systems changes depending on the different walking parameters, analyzing also the data from all the available sensors. Limitations and future works are also presented in this chapter.

Acknowledgements

First of all, to Prof. Atsuo Takanishi, thank you very much for your supervision and wise advice, but above all, for accepting me in your laboratory and giving me the opportunity to get fully involved in the humanoid robotics world and be part of the cutting edge of this field, which has been one of my biggest dreams and which I am now seeing fulfilled.
To Prof. Mitsuo Umezu, Prof. Jun Ohya, Assoc. Prof. Hiroyuki Ishii and Assoc. Prof. Sarah Cosentino, thank you for all your invalulable insights and comments, all of which were key for the constant improvement and successful completion of this work.
To Dr. Kenji Hashimoto, thank you for accepting me in the Biped Group, for all your guidance and your patience and advice for the development of my research.
To Dr. Martim Brandão, thank you very much for all the fruitful discussions about my research. I believe that this work was possible thanks to many ideas from those conversations, which had a great impact on the direction and development of this work.
To Dr. Xiao Sun, thank you very much for all the support and advice during all the graduation process, which can be overwhelming and very confusing. Your guidance was key for a smooth completion of this journey.
To Mr. Di Zhang, thank you for your friendship and for acompanying me through this journey. Having someone on the same quest was of great help for a successful finish.
To Dr. Przemysław Kryczka, thank you very much for all your support and friendship, which helped me a lot both in Japan and again in Italy during my internship at the Italian Institute of Technology. Thanks for all your advices and support with everything.
To Dr. Antonio Morell, thank you very much for your help with the ideas and experiments which were an important part of this work. Also thanks to all the people in the Italian Institute of Technology who helped me during my internship there; the results obtained there were a very important step towards the progress of this work.
I would also like to thank Dr. Aiman Omer for all your good advice, Dr. Matthieu Destephe, for your continuous support and help, and Dr. Gabriele Trovato for Jetlag.
To all the Biped Group members who accompanied me during this quest, especially Dr. Tatsuhiro Kishi, Dr. Takuya Otani, Mr. Takashi Matsuzawa, Mr. Shunsuke Kimura, Mr. Takanobu Matsubara, Mr. Hiroki Mineshita, Mr. Daifeng Wang, thank you all for making me feel as a member of this group, for your support and help with experiments, and for making this quest much more interesting and valuable. Also, thanks to all the members of the Takanishi Laboratory for your support and friendship, especially Mr. Katsuaki Tanaka and Mr. Yihsiang Ma, and I am very thankful and indebted to all the laboratory staff and secretaries for their support during my stay in the Takanishi Lab.
I would also like to express my gratitude to all the members of the Leading Program, as well as all the administrative staff, without whose support none of this would have been possible.

Gracias a Alexis Holgado, Gonzalo Aguirre y Javier Álvarez por darme mi dosis de latinidad y mexicaneidad en este país tan distante del nuestro.
I would like to acknnowledge the Nippon Foundation and the Graduate Program for Embodiment Informatics of the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT, Japan), for their economical support.

And finally, I would like to thank all my family and my friends, for everything, for always being there, for taking care of me and making me the person who I am now.

Yukitoshi Minami Shiguematsu

July 2019

Contents

List of Figures

List of Tables

Abbreviations

A/D	**A**nalog to **D**igital
ATE	**A**bsolute **T**rajectory **E**rror
CAD	**C**omputer **A**ided **D**esign
CoM/CoG	**C**enter **o**f **M**ass/**G**ravity
CPU	**C**entral **P**rocessing **U**nit
D/A	**D**igital to **A**nalog
DC	**D**irect **C**urrent
DoF	**D**egree(s) **o**f **F**reedom
EKF	**E**xtended **K**alman **F**ilter
F/T	**F**orce **T**orque
FFT	**F**ast **F**ourier **T**ransform
IMU	**I**nertial **M**easurement **U**nit
LIP	**L**inear **I**nverted **P**endulum
OKR	**O**pto-**K**inetic **R**eflex
PCI	**P**eripheral **C**omponent **I**nterconnect
PIO	**P**arallel **I**nput **O**utput
PWM	**P**ulse **W**idth **M**odulation
RPE	**R**elative **P**osition **E**rror
SLAM	**S**imultaneous **L**ocalization **A**nd **M**apping
TTL	**T**ransistor-**T**ransistor **L**ogic
VCR	**V**estibulo-**C**ollic **R**eflex
VO	**V**isual **O**dometry
VOR	**V**estibulo-**O**cular **R**eflex
ZMP	**Z**ero **M**oment **P**oint

Chapter 1

Introduction

1.1 Background

Humans are very versatile creatures that have lived on the planet Earth for more than 200,000 years. Since then, they have grown and advanced incredibly as a species. However, we still don't completely understand ourselves, and there are still many mysteries to solve about humans' inner workings. Imitation has always been a powerful method for learning, and in that process, for understanding how things work. Biomimetics, for instance, refers to designs or adaptations inspired by nature, or as the word itself expresses, it means mimicking biology [1–3]. With that in mind, building a machine trying to imitate a human not only on the appearance, but also on how it moves or on how it functions could be useful for a better understanding of them. This has been one of the purposes for building humanoid robots, as in order to develop them we need that deeper understanding of humans, and in turn building them could lead to new insights to further comprehend ourselves.

On the other hand, with the accelerated and active developing of technology and sciences in the last decades, we now live in a world where we can build robots that not only take care of a lot of the work we used to do for ourselves, but also do it in an autonomous way and even are capable of adapting to new situations just by being trained for them or by learning for themselves. As time goes on, they will become more significant in our daily lives, and the possibility of having robots helping us on our daily routines, and not only in specialized places like factories or laboratories is becoming a reality.

Therefore, it is important to design robots with which we could easily interact and that can move freely and without any problems in human environments. Humanoid robots are very good candidates for that, as it is possible to make the interactions with them to be very close to those we have between ourselves. From the point of view of moving in human environments, it is a fact that we built these environments for us to move as comfortably and securely as possible, i.e., they are optimized for our physical structure. This means that these environments are somehow adapted to bipedal walking, and hence, biped humanoid robots seem to be a good direction to take for the design of robots helping us in our daily lives.

However, there are still many problems to solve. One of them is mobility, where bipedal locomotion is one of the best options for human habitats, as the great majority of the environments we live in are designed to maximize the usefulness of humans' physical configuration. And most importantly, as we live in a dynamic environment that can change unexpectedly, it is crucial to have stable, adaptive and robust locomotion in a variety of environments and situations such as uneven terrains, avoiding obstacles, external perturbations, etc.

One of the reasons humans achieve such a robust and adaptable walking is their ego-motion estimation, i.e., the ability to localize themselves in space and to calculate their traveling trajectories w.r.t. their environment. With this information, humans are able to dynamically (re)plan their walking paths, as well as the walking motions themselves, to be able to cope with various situations. In the following sections, we will present a brief summary of the sensory systems used for the ego-motion estimation.

1.2 Human Sensory System

As one of the many goals of movement is to reach or avoid objects or locations in space, information about these objects and locations is very important. However, information about one's own movement, also called self-motion or ego-motion is of utmost importance.

From a developmental point of view, humans first learn to balance before acquiring whole-body motions such as locomotion. Additionally, spatial orientation abilities are

also acquired early as they do not require balance or locomotion, and are later useful for navigation and spatial recall. [4]

For all the above mentioned skills humans are required to select and correctly integrate the input information from a variety of senses. Particularly for balance and locomotion, mainly three inputs are used: visual, gravitational/inertial and proprioceptive. We will now briefly explain each of them.

1.2.1 Visual Inputs

Vision is one of the so called exteroceptive senses, as it provides information relative to external stimuli. Vision relies on photoreceptors in the retina of each eye to generate images of visible light. Humans have two types of photoreceptos: rods, which are very sensitive to light intensity, and cones, which are responsible for color sensing (Fig. 1.1).

To estimate traveled distances from vision, the key visual information is the optic flow, which results from our relative motions with respect to the environment [5].

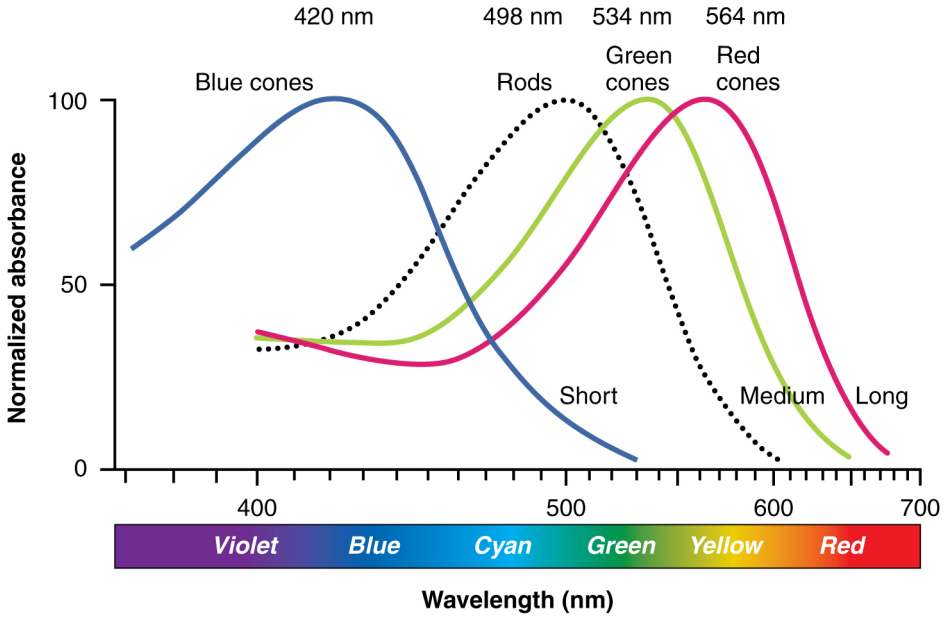


FIGURE 1.1: Color sensitivity of the human eye. Wavelength in the horizontal axis and normalized absorbance in the vertical axis. [6]

1.2.2 Gravitational/Inertial Inputs

For the gravitational and inertial cues, humans rely on two sources, the vestibular system, located in our inner ears on our head, and on cues from the body, called somatic graviceptors.

Vestibular System

The vestibular system comprises the semicircular canals and the otoliths inside our inner ears. There are three semicircular canals in each labyrinth, which are orthogonal to each other, and are responsible for the sensing of rotational movements of the head. Particularly, semicircular canals sense angular velocities of the head. Otoliths, on the other hand, are responsible for sensing linear accelerations of the head. Humans have two otolithic organs on each side: the ultricle, which senses accelerations and head tilts on the horizontal plane, and the saccule, which senses accelerations and head tilts on the vertical plane.

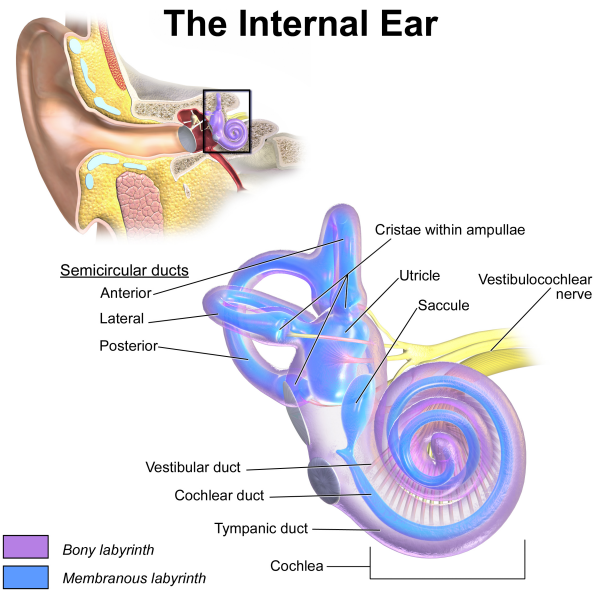


FIGURE 1.2: Structure of the internal ear. (*Wikijournal of Medicine, 2014. Internet resource.*)

Somatic Graviception

However, there are studies showing that we also "feel" gravity or linear accelerations with our body, a sense called **Somatic Graviception**. [7] It is theorized that there are

two somato-graviceptive sources, one being the kidneys and the second being the blood inside large vessels on the body. Presumably, the mass of the blood exerts inertial forces on the ligaments which support the vessels against the gravitational load.

1.2.3 Proprioceptive Inputs

Proprioception is the sense of the relative position and movements of parts of our own body, as well as the strength of effort used by them for moving. In humans, there are many sources of this information.

Mechanoreceptors

One of them are the mechanoreceptors, which are sensory neurons or peripheral afferents located within joint capsular tissues, ligaments, tendons, muscle, and skin. These sense deformation or stimulation of the tissues and there are classified depending on the type of tissue they are located on.

Those located in the joints[8–10] are classified in four primary types:

Type I : Typically located in the superficial layers of the joint capsule. Physiologically, these receptors are low-threshold, slowly adapting, and they signal static joint position, changes in intraarticular pressure, and the direction, amplitude, and velocity of joint movements.

Type II : Present in the fibrous capsules of all joints but are reported to be present in greater number in distal joints than in proximal joints. They are low-threshold, rapidly adapting receptors which signal joint acceleration and deceleration during both active and passive joint movements.

Type III : Primarily confined to the joint ligamentous structures. They are high-threshold, slowly adapting structures and become active or stimulated only toward the extreme ranges of joint motion where the ligamentous structures become taut or if extreme joint displacement or joint traction is maintained.

Type IV : Typically distributed throughout the fibrous joint capsule, adjacent periosteum, and articular fat pads. They represents the pain receptor system of articular

tissues and are activated only by marked mechanical deformation or chemical irritation.

On the other hand, those located on the skin are specialized in the perception of pressure, vibration and texture[11, 12] and are also classified in four groups:

Slowly Adapting Type 1 (SA1) Merkel's disks are very sensitive to points, edges and curvature., and have a high spatial resolution, as they can resolve spatial detail of 0.5 mm. They seem to be responsible for form and texture perception.

Slowly Adapting Type 2 (SA2) Ruffini endings are very good at sensing skin stretch, with which they can perceive the direction of motion or force when the skin is stretched, and also can sense the shape and position of parts of our body through the pattern of skin stretch.

Fast Adapting Type 1 (FA1) Meissner's corpuscles are responsible for the detection and discrimination of low frequency vibration, slip between the skin and a held object, and are the best to sense sudden forces acting on held objects. These seem to provide feedback signals for grip control.

Fast Adapting Type 2 (FA2) Pacinian corpuscles are responsible for the sensing of high frequency vibrations.

Proprioceptors

In addition to mechanoreceptors, there are also proprioceptors, which are specialized organs located within joints, muscles and tendons. They are sensitive to both tension and pressure, and relay muscle dynamics information to the central nervous system. They also play an important role on the conscious appreciation of the position of body parts with respect to gravity, also called kinesthetic sense. Most proprioceptive information is processed at a subconscious level, and for this reason we do not have to dedicate any conscious activity for tasks such as maintaining posture or position body parts [13]. They are divided in:

Muscle Spindles Small sensory organs enclosed within a capsule, are found throughout the body of a muscle in parallel with extrafusal muscle fibers. They sense

changes in the length of the muscles through intrafusal muscle fibers, which have a central region wrapped by sensory dendrites of the muscle spindle afferent. They can also sense the speed of these changes, and they contribute to both sense motion and position of limbs. They prevent muscle overstretching and muscle fiber damage [14, 15].

Golgi Tendon Organs As their name indicates, they are located near the origins and insertion of skeletal muscle fibers into the tendons of skeletal muscles. They sense changes in muscle tensions and are also proprioceptive information sources.

1.2.4 Sensor Fusion and Reweighting

All the above mentioned sensors, taken individually, are very imprecise as motion sensors, as they are specifically designed to acquire a particular kind of information which alone would give few or flawed clues about posture or self-motion. For instance, the vestibular system presents two significant problems:

Rotation Problem As the semicircular canals are fixed inside the head, angular velocity is coded in a head centered reference frame, which somehow has to be transformed to obtain information about the body.

Linear Acceleration Problem From Einstein's equivalence problem, the otoliths cannot distinguish between linear accelerations and body tilting when in a gravitational field.

Then, for instance, one solution for the Rotation Problem would be to have information about the head motion relative to the world.

Sensor Fusion is very important, as it can combine the information from various sensors' (of different modalities)redundant data to extract more precise and useful information than that from the sensors individually.

There are many models trying to explain how humans perform sensor fusion for egomotion.

For this research, we decided to employ the model proposed by Clemens et al. [16] (Fig. 1.3), because one of the interests is to take Somatic Graviception into account.

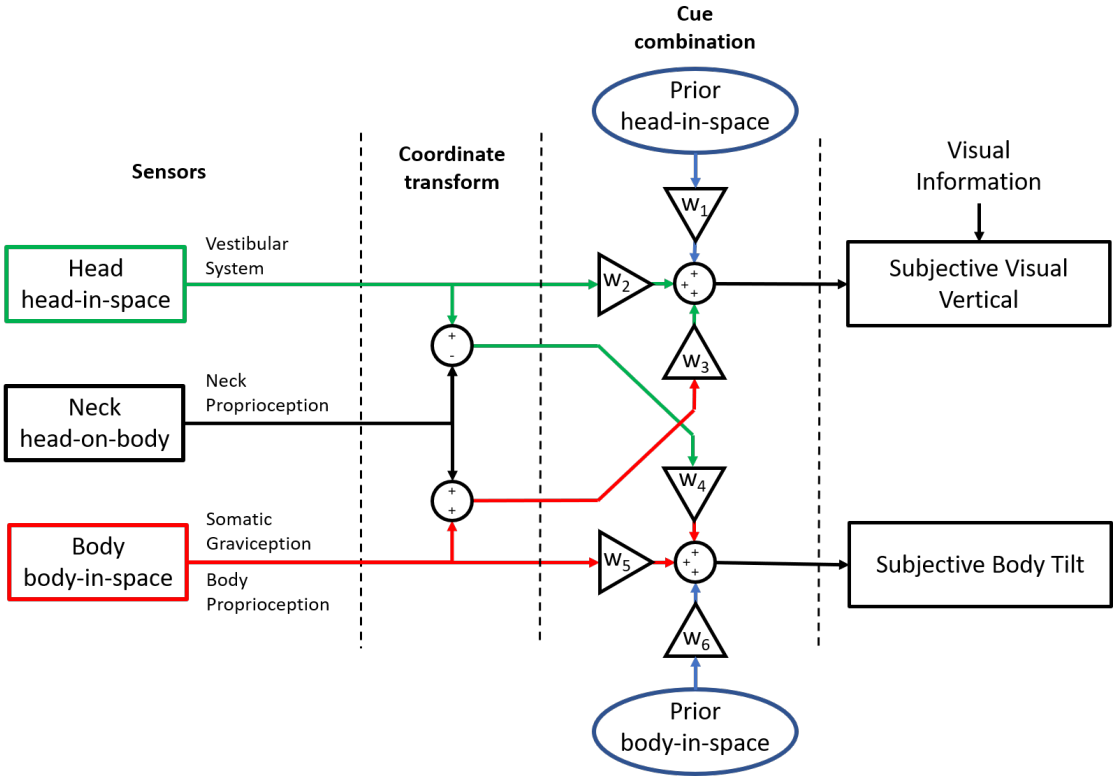


FIGURE 1.3: Indirect Pathway Model diagram, modified from [16]. The different weights applied to each input are denoted by w.

Moreover, humans not only combine different sensor's data, but also change the weight they put on each sensory input depending on the situation [17]. Also, both visual and vestibular systems are located on the head, and the manner in which they are placed is very particular, but the reason for this is not completely understood.

1.2.5 Human Navigation

Navigation refers to the monitoring and controlling of the motion of a moving agent, or in our case, of a human being. It involves the planning of the motion through an environment, or in the case of becoming lost, reorienting and reestablishing the motion towards the destination. One fundamental element of navigation is updating position and orientation while moving. Updating methods can be classified as follows according to the type of information used [18]:

1. Position based navigation, also called *pilotage* or *piloting*, relies on external signals such as visual, auditory or odorous landmarks, which indicate the observer's current position and orientation.

2. Velocity based navigation, also called *dead reckoning* or *path integration*, relies on the sensing of the traveler's instantaneous speed and direction of motion, from which linear and rotational displacements w.r.t. the initial position are calculated by integration. Therefore, it requires no map. The measurements of velocity can be allothetic (external) signals as optic or acoustic flow, or idiothetic (internal) signals as those from the proprioceptive or the vestibular systems.

3. Acceleration based navigation, also called *inertial navigation* or also *path integration*, requires information about linear and angular accelerations, to then integrate it twice to obtain position and orientation. In humans, this information can be obtained from the vestibular system, and also from somatosensory signals, such as somatic graviception.

For this work, we focus on path integration, i.e., the updating of position and orientation from velocity and acceleration information, as it is said to allow one to venture into unfamiliar territory for the purposes of seeking a destination, and as one explores an unfamiliar region of space, it provides the traveler with an ongoing estimate of current position, allowing the traveler to gradually integrate the isolated perspective views encountered into an internal representation (cognitive map [19]) suitable for subsequent piloting. [20]

Also, there is a great body of research proving that humans are able to accomplish path integration using only their own senses, be it only from vision and optic flow [21, 22], using proprioception, gravito-inertial cues, or a combination of all these.

Moreover, some studies show age-related differences in path integration abilities, such as an aging-related deficit to obtain body rotation information from the vestibular system, or age-related differences related to cognitive sources as speed of processing and working memory capability [23].

1.3 Related Research

Robots have been used to understand humans, analyzing them from many different points of view.

Computational Brain (CB)[24] was developed for exploring the processing of the human brain in dealing with the external environment through the use of a biped humanoid robot (Fig. 1.4). It has 50 DoF, and both its mechanical configuration and sensing system were designed to be as close as possible to humans, to also bring its performance closer to us.

They implemented various humanlike ocular-motor responses such as vergence, saccades, smooth pursuit and the vestibulo-ocular reflex (VOR) [25].

FIGURE 1.4: Humanoid robot CB [24].

HRP-2 [26] is a humanoid robot designed to be light and compact, with the ability to perform tasks as cooperative works in the open air (Fig. 1.5). It has 30 DoF, and the motion range of each joint was set to be as that of humans in order for the robot to be able to perform human tasks as well as human beings.

Using HRP-2, in [27] a control method for reaching motions based on a neuromuscular system defined by a pair of antagonist muscles based on neurobiological motor principles were developed and tested in simulation.

Whether the central nervous system of primates uses a body-centered or an eye-centered frame for reaching motions was explored in simulation in [28]. They used a simulation

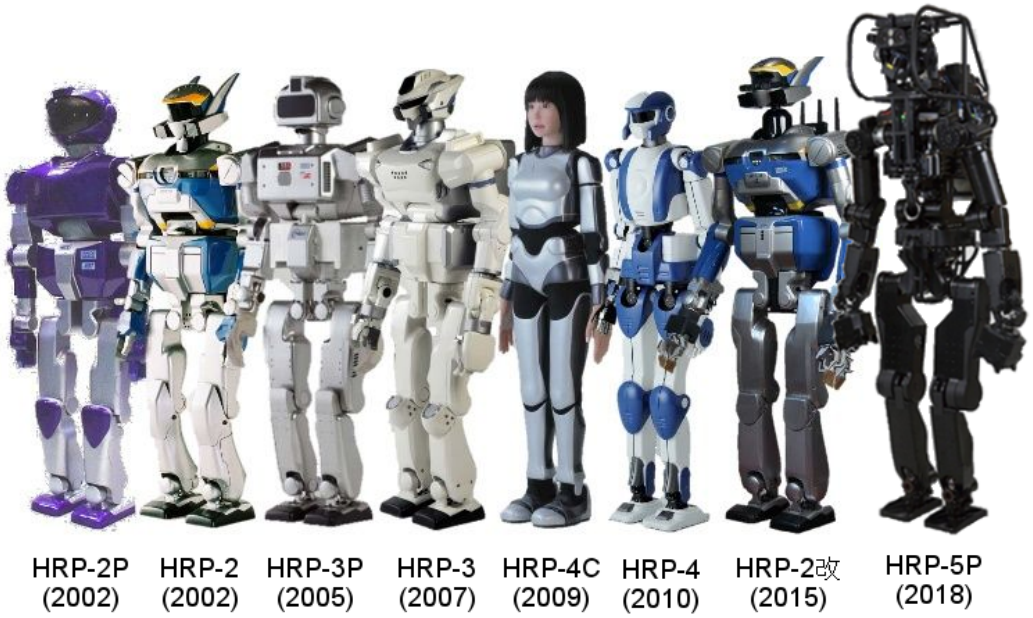


FIGURE 1.5: The HRP Humanoid robot series. (modified from [30])

of the robotic platform HRP-2, showing that for the robot, the eye-centered frame was more robust with respect to perturbations as proprioceptive biases and sensory delays.

In [29] a robot steering method by moving the head joint was developed, inspired by the fact that in human locomotion the head plays an important role in guiding and planning locomotion.

KOBIAN [31] is a biped humanoid robot developed for whole body emotion expression (Fig. 1.6). It is 1470 mm tall, weights 62 kg and has 64 DoF, of which 27 are on the head for facial expressions.

In [32], a head stabilization controller based on information from an IMU monted on the robot's head. This stabilization is based on the angular Vestibulo-Collic Reflex (aVCR) and the linear Vestibulo-Collic Reflex (lVCR) from humans, which are reflexes that compensate the motions induced by walking to stabilize the head orientation and the gaze when fixated on a point in close proximity [33–35]. The model used for the robot stabilization controller was based on the study in [36].

On the same line, in [37] the above mentioned aVCR and lVCR, along with the VOR and Opto-Kinetic Reflex were also implemented on simulation on iCub and on the humanoid robot SABIAN.

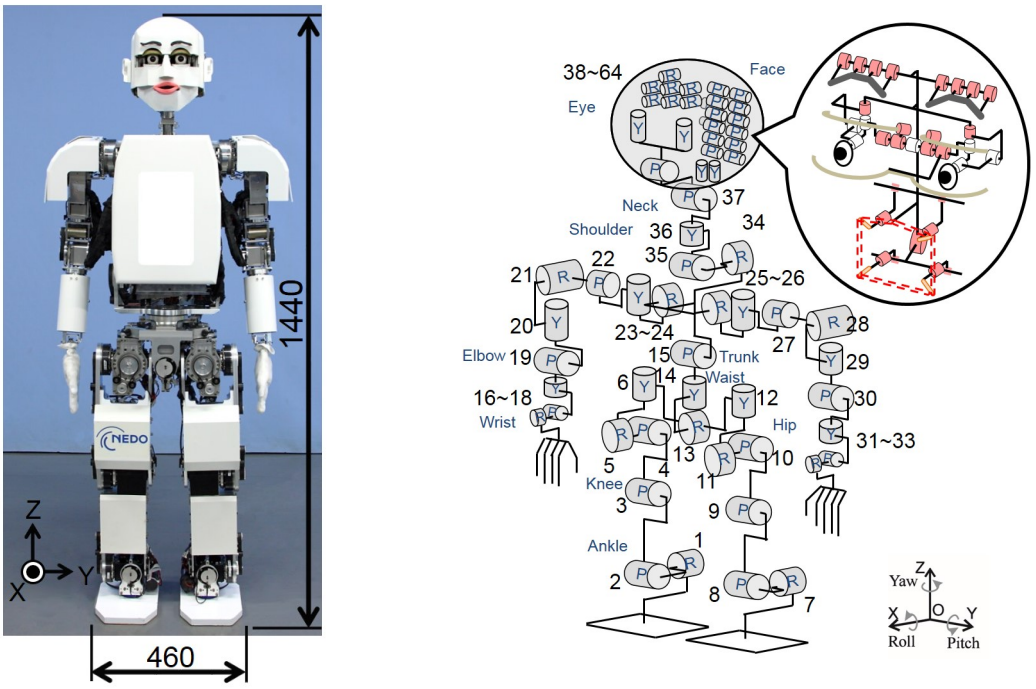


FIGURE 1.6: Robotic platform KOBIAN-RIV (left) and DoF configuration (right).

The humanoid robot iCub was developed with the objective to replicate the physical and cognitive abilities of a 2.5-year-old baby [38, 39]. It is 90 cm tall, weights 23 kg, and has 53 DoF. In [40], the Vestibulo-Ocular Reflex (VOR) was studied with this robot. They implemented and compared two bio-inspired models of the VOR on iCub, a Feedback Error Learning (FEL) model, and a decorrelation model. The FEL model was used to investigate the cooperation between the Opto-Kinetic Reflex (OKR) and the VOR, while the decorrelation model was used to investigate the structure and plasticity of the cerebellar cortex aiming to better understand the cerebellar microcircuit. There are also works for gaze stabilization with this platform [41].

Posturob is a biped humanoid 'stance control' robot, built to study human posture [42]. It uses artificial pneumatic muscles as actuators, with springs as tendons and force sensors at the muscle-tendon fixations, mimicking Golgi tendon organs. For the sensory system, they used a rotatory potentiometer fixed to the robot's ankle joint axis as joint proprioception, compression load cells under the forefeet and heels for force sensing, gyroscopes as the semicircular canals, and accelerometers as otoliths. With this robot, biologically inspired models of human posture were developed and tested. In [43] they focused on postural disturbances as external forces and motions of the support surface on which the body stands. In [44], they evaluated a posture control concept

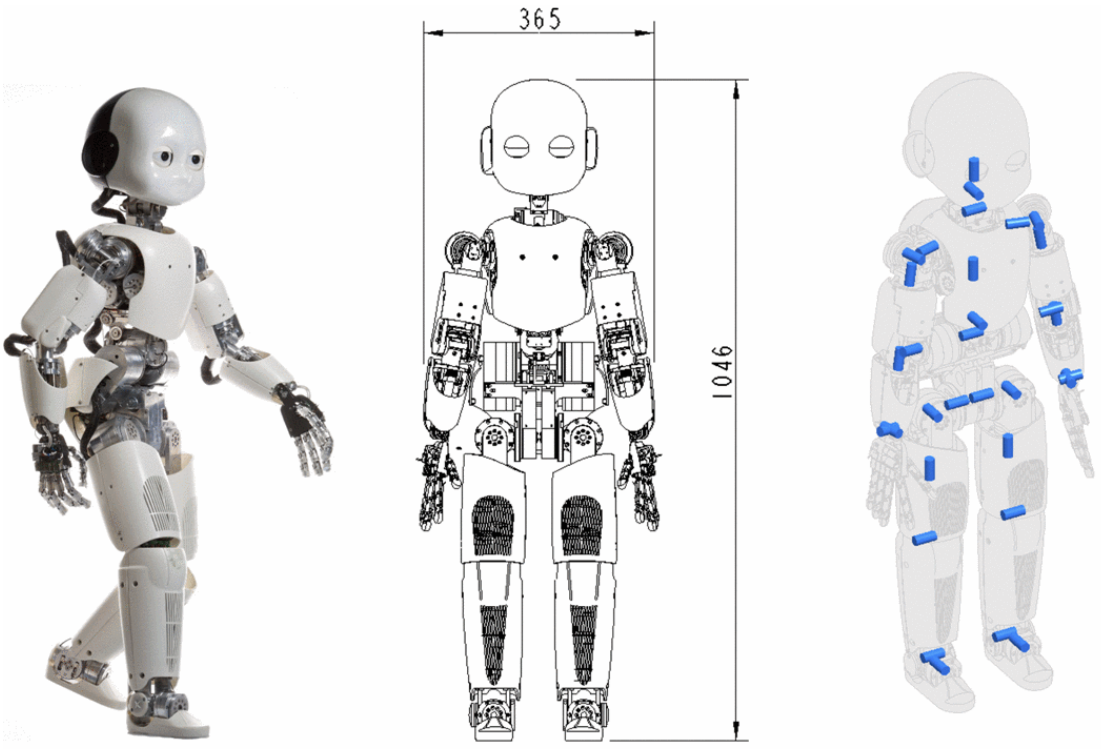


FIGURE 1.7: The iCub platform (left), approximate dimensions (center) and DoF configuration (right). Taken from [39].

from neuroscience called disturbance estimation and compensation, which differs from typical state estimation systems by not including a dynamic model of the body.

TABLE 1.1: Other robots

Robot		**HRP-2 [26]**	**CBi [24]**	**Posturob II**
Sensor	**Vision**	Three cameras	Foveated cameras	None
	Graviception	Gyroscope (torso)	IMU (head & hip)	Gyro & Accelerometer (head)
	Force/Torque	Feet	Feet and main joints	Feet
	Joints	Encoders	Encoders	Goniometers
Robot Design		Human based	Human based	Study based (human posture)
Sensing System		Task based (various)	Human based	Study based (human posture)

FIGURE 1.8: Humanoid robot Posturob II.

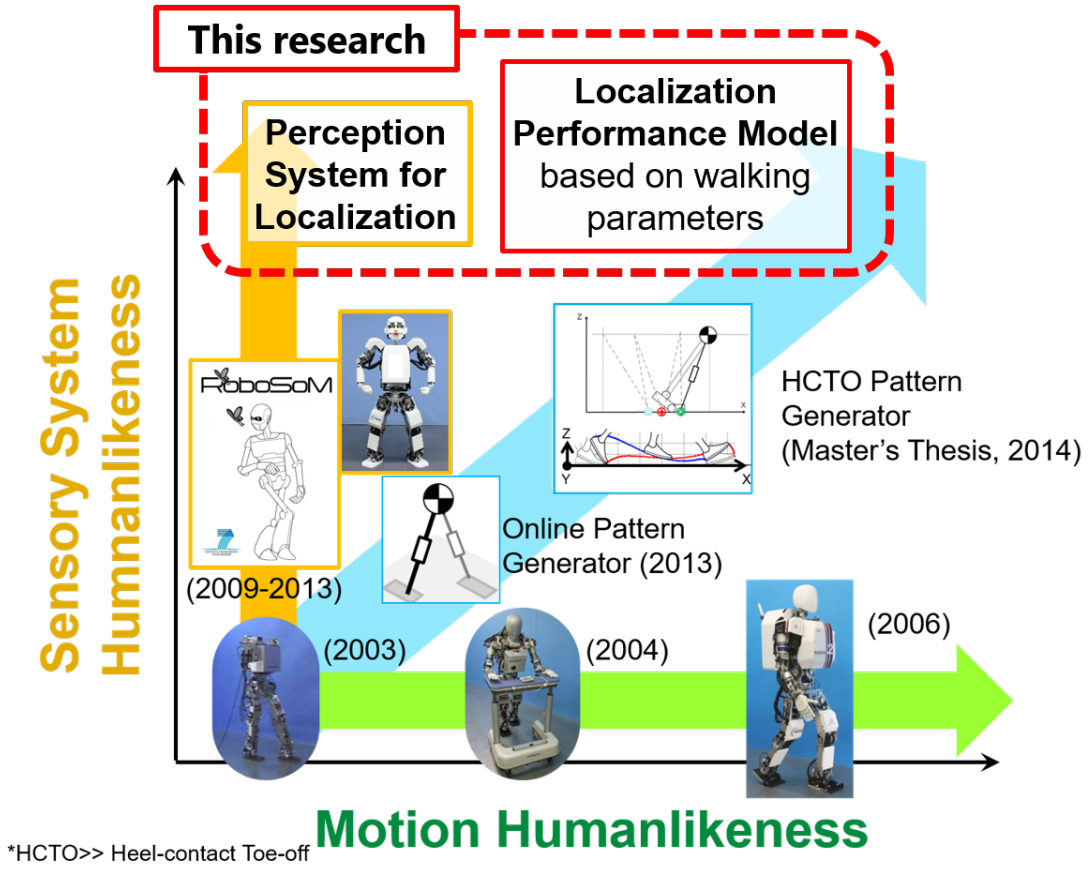


FIGURE 1.9: The roadmap of this thesis

1.4 Objectives and contributions

As research about the human walking have already been thoroughly done, we can now aim to the understanding of the human sensory system, beginning from how each sensing mechanism is located on the body, what kind of information they obtain, and how we humans combine and use that information to realize all the different motions we are capable of, focusing in the action of walking for the present research. Moreover, we are interested in the ego-motion estimation, and on assessing how the walking motion affects the performance of this estimation.

The objectives for the present work are the following:

1. Development of a Perception System for Localization

   (a) Based on the Human Sensory System, i.e., using visual, proprioceptive and gravitational/inertial inputs

   (b) Taking into account both Somatic Graviception and the Vestibular System as gravitational/inertial inputs

   (c) Separating the localization problem into head and body localization

2. Confirm the effects of Walking Parameters on the Localization Performance of the Perception System using Robotic Science

   (a) Step length (walking speed) effects

   (b) Walking style and symmetry effects

1.5 Thesis Outline

The presented thesis is constituted by six chapters which will be briefly summarized below, and whose structure is presented in Fig. 1.10.

Chapter 1 introduces the research background. More specifically, it presents how humans perceive their motion, briefly describing the different systems participating in it, particularly focusing on the Indirect path model for self-motion estimation, which proposes that humans estimate the motion of their head and their body separately, and then combine this information through the kinematics of the neck. Moreover, it contains the objective of this research and the comparison with other related researches in this field.

Chapter 2 introduces the sensor system proposed for the head motion estimation of the robot, as well as the algorithm to be used to fuse the different sensor inputs. The estimation system comprises visual and gravito-inertial inputs, obtained from a camera and an IMU on the robot's head respectively. The results from the simulation tests, as well as those from the first prototype are presented.

Chapter 3 presents how the walking motions affects the localization performance of the head localization system proposed in the previous chapter. Particularly, step length is the modified walking parameter. Three VO algorithms are compared: a direct, a semi-direct and an indirect VO algorithm. The performance is measured through the absolute trajectory error, and the relative pose error.

Chapter 4 introduces the sensor system proposed for the body motion estimation of the robot, as well as the algorithm to be used to fuse the different sensor inputs for the estimation. The estimation system comprises proprioceptive and gravito-inertial inputs, obtained from joint encoders and force-torque sensors on the feet, as well an IMU on the robot's trunk, respectively. Results of the preliminary experiments with the biped humanoid robot WALKMAN from the IIT are presented.

Chapter 5 presents how the walking motions affects the localization performance of the body localization system proposed in the previous chapter. The modified parameters are step length, walking style and walking symmetry. The performance is measured through the absolute trajectory error, and the relative pose error.

Chapter 6 discusses the quantitative and qualitative results of this work, analyzing how the localization performance of the proposed systems changes depending on the different walking parameters, analyzing also the data from all the available sensors. Limitations and future works are also presented in this chapter.

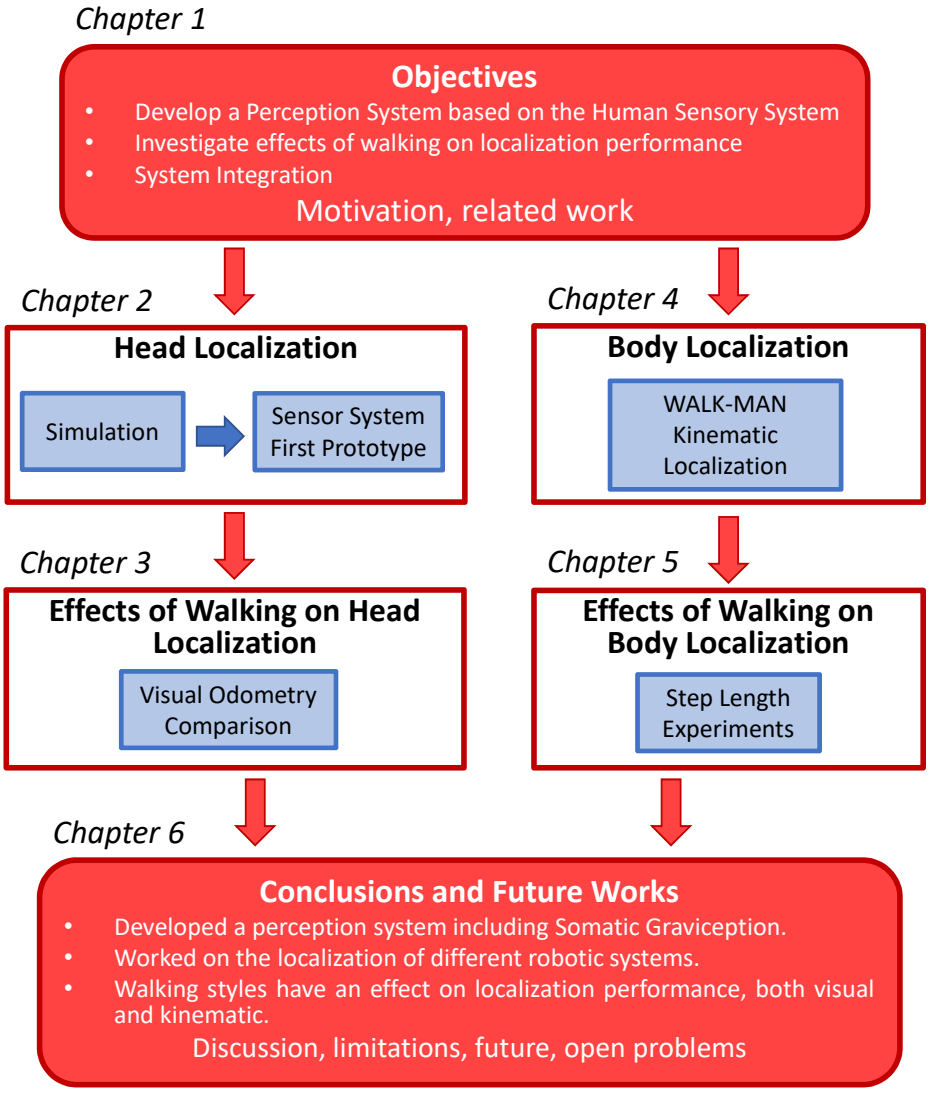


FIGURE 1.10: Structure of this thesis

Chapter 2

Head Localization

2.1 Introduction

In Chapter 1 we presented the senses humans use in order to estimate their ego-motion, to then combine the information from all those senses and finally to localize themselves in space. One suggested model for how humans combine the information from different sources is the indirect path model [16]. This model proposes the separation of the body motion estimation in two different estimations, one for the position and orientation of the head, and one for the position and orientation of the body. In this chapter we will present the development of the Head Localization System, the system for the estimation of the position and orientation of the humanoid robot's head.

First we will present the possible sensors to use, in order to emulate the same inputs that humans use for this task. After that, we will discuss about the different methods to combine the information from the sensors to obtain an estimation of the position and orientation of the system, to afterwards use that information for the robot's localization in space. With that in mind, the objectives for this chapter are as follows:

a) Design a sensor system with similar inputs as those used bu humans for the head position and orientation estimation

   • Choose the suitable sensors for the system, to obtain similar information as that from human senses

- Select the fusion algorithm to combine the information from the chosen sensors

b) Evaluate the proposed system feasibility in simulation

c) Build and evaluate a real prototype of the proposed system

To achieve this, in the following section we will present the proposed system, introducing the chosen sensors as well as the selected fusion algorithm. Tests of the system in a simulation environment are shown, and finally the real prototype is presented, also evaluating its performance through simple experiments.

2.2 Sensors

According to the indirect path model [16], the required information to estimate the position and orientation of the head is the images obtained from the eyes, the position of the eyes inside the head, and the inertial and gravitational information from the vestibular system.

To emulate the different sensory inputs, we chose sensors that would provide us with similar data. Therefore, for the visual input we chose cameras for the vestibular inputs we chose Inertial Measurement Units (IMU).

2.3 Visuo-Inertial Fusion

For the fusion of inertial and visual information, we decided to use an Extended Kalman Filter (EKF) based approach, as there are some clues pointing to the fact that humans employ probabilistic models during sensorimotor learning [45]. The EKF is a nonlinear and recursive Bayesian estimator that tries to estimate of unknown variables of a system, using a series of observed noisy and inaccurate measurements. It does this by estimating a joint probability distribution over the variables for each time frame. The algorithm has two steps, prediction and update. In the prediction step, the filter estimates the value of the variables from the past information and a model of the phenomenon. Then, in the update step, the filter uses observed measurements usually from sensors to correct the

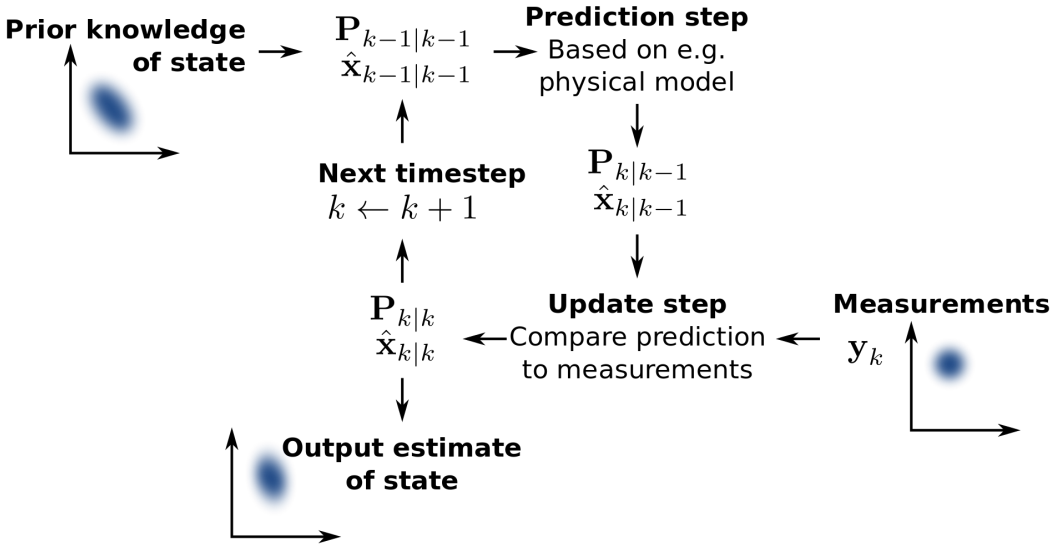


FIGURE 2.1: Kalman filter diagram. The system state is denoted by $\hat{x}$, where $\hat{x}_{k|k-1}$ is the state estimate at time step k before the k-th measurement y_k has been taken into account and $\hat{x}_{k|k}$ is the estimate taking into account the measurement yk. $P_{k|k-1}$ and $P_{k|k}$ are the corresponding uncertainties.

estimated values in the predictions step, through a weighted average where the weights depend on the certainty of the information.

Using EKFs, there are two main approaches to recover the motion of the vision sensor fusing visual and inertial information: *Loosely Coupled* and *Tigthly Coupled* [46]. In the Tightly Coupled approach, a single high-order estimation filter is used, i.e., the estimated states include not only those of the visuo-inertial system, but also those of the tracked features in the visual thread, and even unknown camera intrinsic parameters and inertial sensor bias and scale parameters. This large number of states makes this approach very computationally heavy and thus slow and difficult to tune due to the large amount of parameters. Also, the implementation becomes complicated due to the difference of the sampling rate of the inertial and visual observations.

On the other hand, in the Loosely Coupled approach the inertial thread and the visual odometry thread run independently, exchanging information to correct their estimates. Usually the inertial thread is used for the prediction phase, whereas the visual odometry thread is used to correct for integration errors from the inertial estimation. This makes the loosely coupled approach less computationally heavy, as well as more versatile in order to use and combine information from other sensors.

Loosely coupled visuo-inertial systems can be considered as virtual stabilized cameras, with the advantage of having no moving parts and superior dynamics compared to mechanically stabilized cameras [46]. Primates, humans included, have a mechanically stabilized visual system, where the image on the retina of the eye is stabilized through reflexes controlled by the vestibular system, such as the vestibulo-ocular reflex (VOR) for high frequency motions, or the optokinetic tracking reflexes for low frequency motions.

From the above, we decided to use a loosely coupled approach for our purpose because of its similarity with the human sensory system, as well as the ease and versatility of its implementation. Especifically, we used the method proposed in [47]. They present a sensor fusion framework able to process delayed, absolute and relative measurements from multiple sensors, while allowing self-calibration of the sensor system. We will describe the framework in the following paragraphs.

This framework is based on an iterated EKF, where the state predictions are driven by IMU measurements. There is a core set of states, and depending on the additional sensors, states can be added to handle those additional measurements. The core states

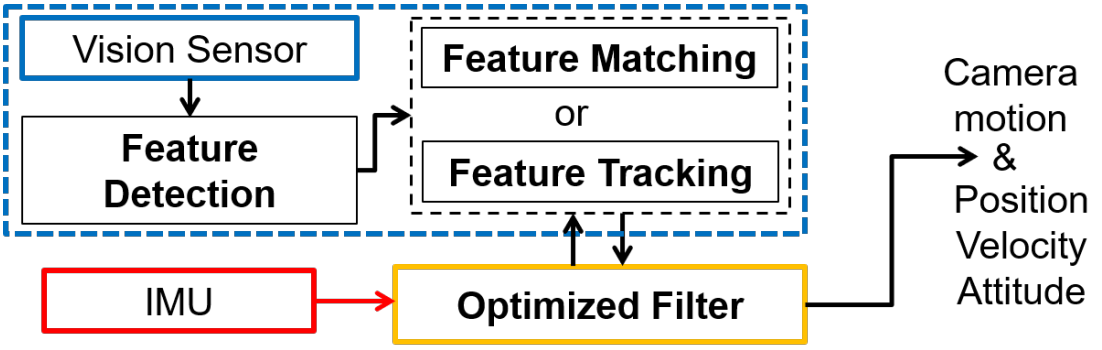


FIGURE 2.2: Tightly coupled fusion system diagram.

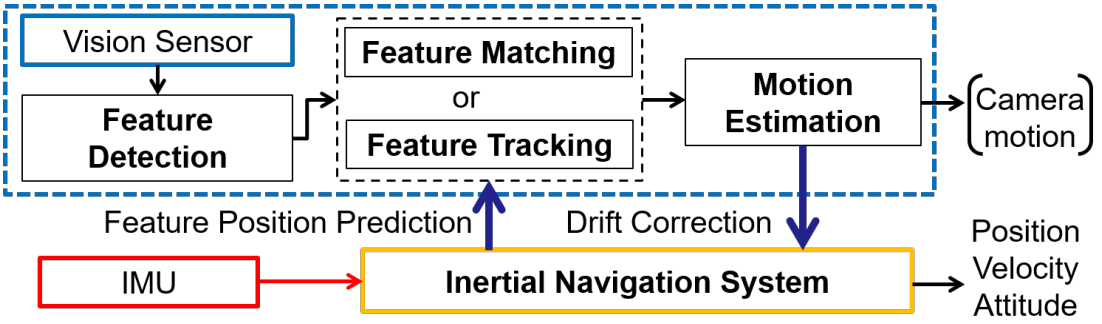


FIGURE 2.3: Loosely coupled fusion system diagram.

can be expressed as:

$$x^T = \begin{bmatrix} {p_W^i}^T & {v_W^i}^T & {q_W^i}^T & {b_\omega}^T & {b_a}^T \end{bmatrix} \tag{2.1}$$

where p_W^i, v_W^i and q_W^i correspond to the relative position, velocity and attitude (expressed as a quaternion) of the IMU w.r.t. the world frame, expressed in the world frame. The remaining b_ω and b_a correspond to the gyroscope and acceleration biases of the IMU.

For the monocular visuo-inertial case, the added states are:

$$x_{vis}^T = \begin{bmatrix} \lambda & q_i^{cT} & p_v^{WT} & q_v^{WT} \end{bmatrix} \tag{2.2}$$

where λ is the estimated visual scale, q_i^c is the rotation transformation between the camera and the IMU, and p_v^W and q_v^W are the position and attitude drifts of the visual odometry system w.r.t. the world frame.

Measurements from an IMU are used for the propagation of the states, as already mentioned, following the equations below:

$$\dot{p}_W^i = v_W^i \tag{2.3}$$

$$\dot{v}_W^i = {C_{(q_W^i)}}^T (a_m - b_a - n_a) - g \tag{2.4}$$

$$\dot{q}_W^i = \frac{1}{2} \Omega (\omega_m - b_\omega - n_\omega) q_W^i \tag{2.5}$$

$$\dot{b}_\omega = n_{b_\omega} \quad \dot{b}_a = n_{b_a} \quad \dot{\lambda} = 0 \tag{2.6}$$

$$\dot{p}_i^c = 0 \quad \dot{q}_i^c = 0 \quad \dot{p}_v^W = 0 \quad \dot{q}_v^W = 0 \tag{2.7}$$

where a_m and ω_m are the measured acceleration and angular velocity from the IMU, g is the gravity vector in the world frame and n_a, n_ω, n_{b_ω} and n_{b_a} are zero-mean white Gaussian noises. $\Omega(\omega)$ is the quaternion-multiplication matrix of vector ω [48], defined as:

$$\Omega(\omega) = \begin{bmatrix} 0 & \omega_z & -\omega_y & \omega_x \\ -\omega_z & 0 & \omega_x & \omega_y \\ \omega_y & -\omega_x & 0 & \omega_z \\ -\omega_x & -\omega_y & -\omega_z & 0 \end{bmatrix} \tag{2.8}$$

The measurements model from the camera pose is defined as:

$$z_{vis} = \begin{bmatrix} p_v^c \\ q_v^c \end{bmatrix} = \begin{bmatrix} C_{(q_v^W)} \left(p_W^i + {C_{(q_W^i)}}^T p_i^c \right) \lambda + p_W^v + n_{p_v} \\ q_i^c \otimes q_W^i \otimes q_v^W \otimes \delta q_{n_{q_v}} \end{bmatrix} \tag{2.9}$$

where $C_{(q)}$ denotes the rotation matrix of quaternion q, n_{p_v} and n_{q_v} are measurement noises for the position and attitude respectively, also modeled as a zero-mean white Gaussian noise, and $\delta q_{n_{q_v}}$ represents the small rotation caused by the attitude noise n_{q_v}. Also, $\otimes$ denotes a quaternion multiplication as defined in [48].

Visual Odometry

One source of information for the visuo-inertial fusion system is an estimate of the position and orientation of the system. This is usually obtained from a visual odometry algorithm. Visual odometry (VO) is the process of estimating the egomotion of an agent (e.g., vehicle, human, and robot) using only the input of a single or multiple cameras attached to it. This is achieved by comparing the sequential images obtained from the sensor and incrementally estimating the pose of the vehicle through examination of the changes that motion induces on the images of its onboard cameras. For VO to work effectively, there should be sufficient illumination in the environment and a static scene with enough texture to allow apparent motion to be extracted. Furthermore, consecutive frames should be captured by ensuring that they have sufficient scene overlap [49, 50]. These algorithms are normally classified based on two parameters, the information they use from the images, and the amount of pixels used from the input images [51]. Depending on the information they use from the images, the algorithms can be:

Direct Pixel intensity information is used directly, making them fast, but prone to errors caused by changes in lighting conditions

Indirect The raw sensor data is pre-processed, extracting features such as corners, edges, or more sophisticated feature descriptors, which makes them more robust to lighting changes, but computationally heavier because of the feature calculation process.

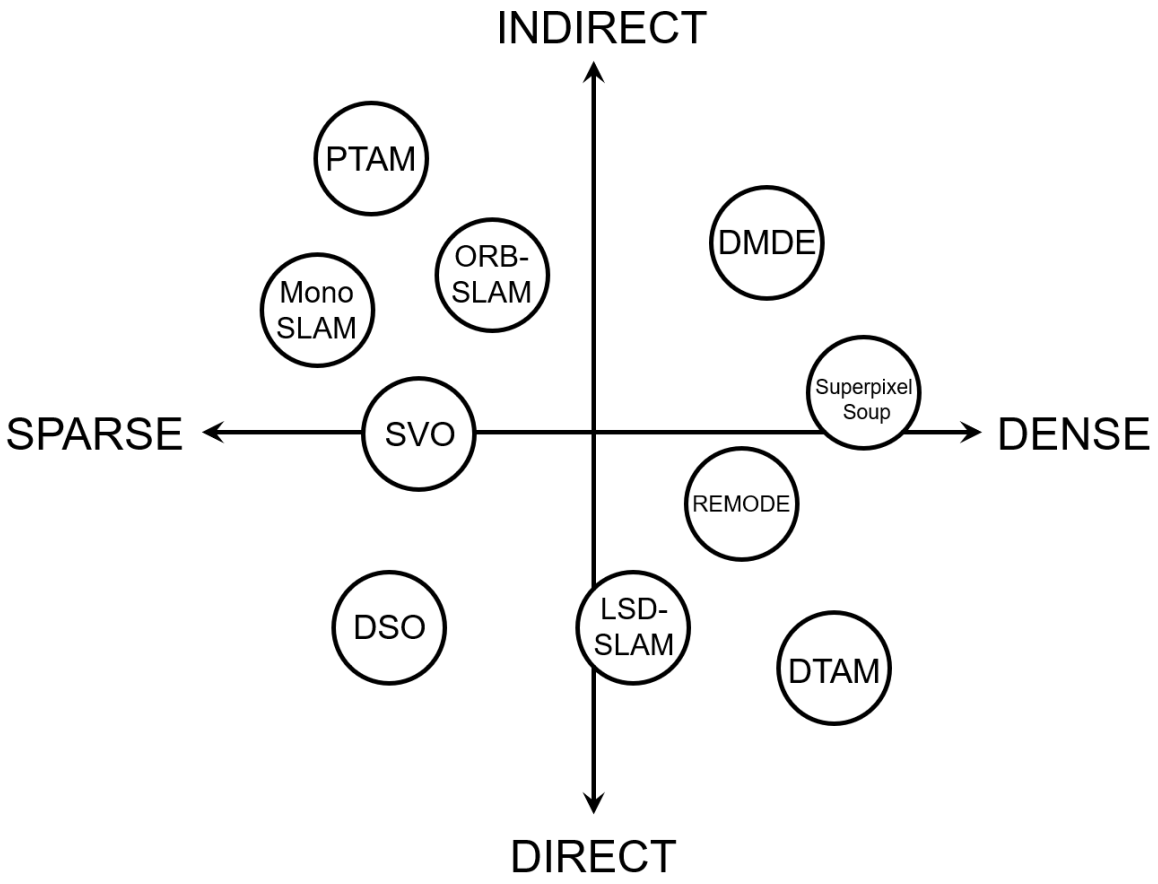


FIGURE 2.4: Classification of different visual odometry algorithms. (Modified from [52])

On the other hand, depending on the amount of pixels used from the input images, the algorithms are classified as *Dense* or *Sparse*.

Sparse These methods use and reconstruct only a selected set of independent points

Dense These methods attempt to use and reconstruct all pixels in the 2D image domain

For this chapter, we decided to test *Sparse* methods, as we are focusing on localization and do not need to reconstruct a map from the visual input.

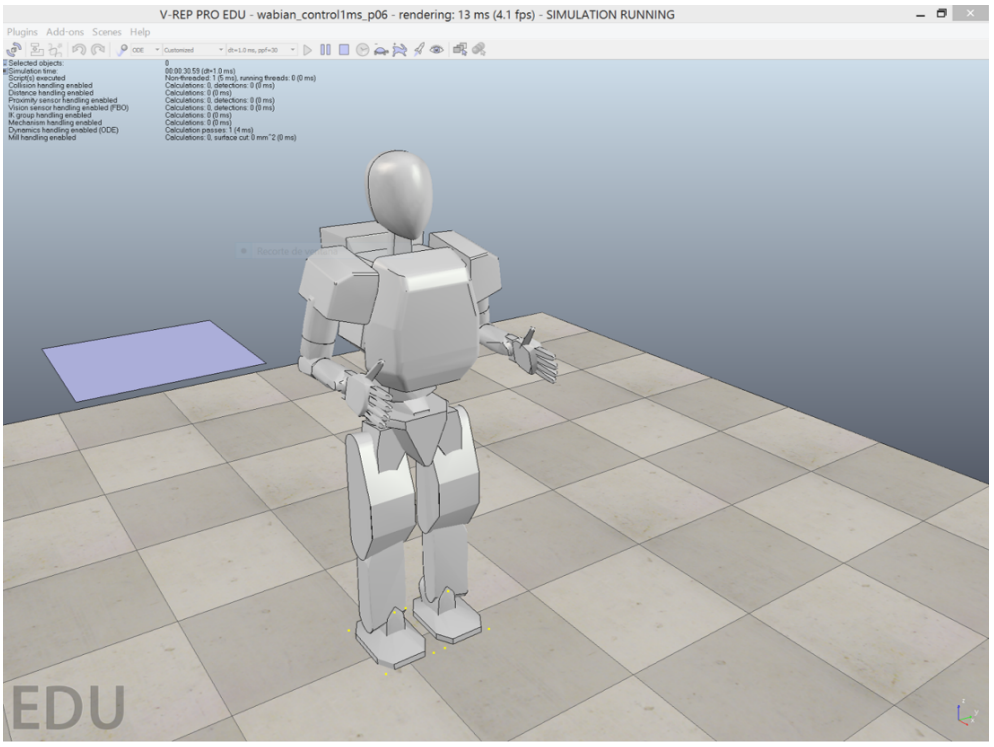


FIGURE 2.5: Simulation environment of the used robot, WABIAN-2R.

2.4 Experiments

2.4.1 Simulation Experiments

For our simulation experiments, we use the humanoid robot WABIAN-2 [53], a human-size humanoid robot, 1.5 meters tall, weighting 64kg and having 41 DOFs. The simulations are performed using the Open Dynamics Engine (ODE) for physics simulation on the V-REP robot simulator [54]. A snapshot of the simulation environment can be seen in Figure 2.5.

The IMU was modeled using MATLAB Simulink © (Fig. 2.6).

Simulations using ROS [55] were also performed.

2.4.2 Visuo-Inertial Prototype

A prototype was built using a monocular camera and an IMU (Fig. 2.9), and a qualitative experiment was performed. For this, the prototype was mounted on a helmet, which was worn by a person. The person performed a closed quadrangular shaped trajectory

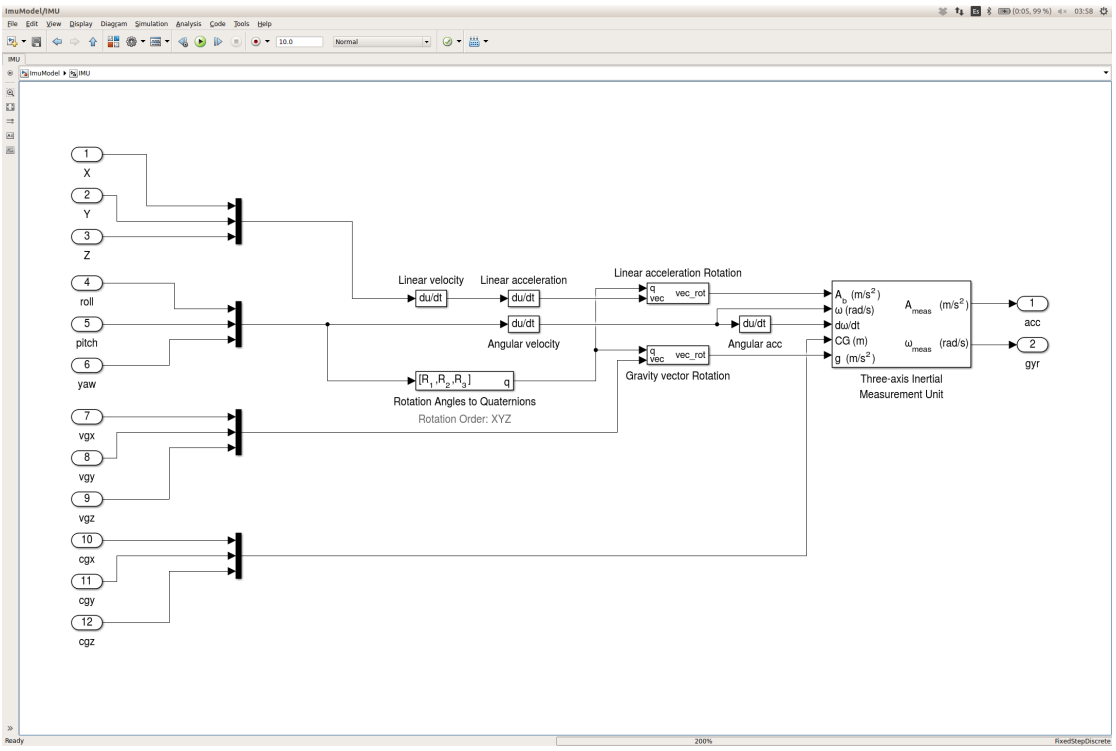


FIGURE 2.6: Used model for the IMU.

(Fig. 2.11), and the path was estimated both using only a sparse and semi-direct VO algorithm [56], and using a visuo-inertial fusion algorithm [47]. A snapshot of the experiment can be seen in Figure 2.10. As can be seen in the same figure, both algorithms overall estimated the trajectory correctly. However, the VO algorithm stopped the estimation in some places, resulting in a discontinuous trajectory, whereas the visuo-inertial fusion algorithm successfully compensated those problematic parts for the VO with the IMU data, resulting in a completely continuous trajectory.

2.5 Discussion

We will organize the discussion of this chapter according to the objectives set in the beginning.

2.5.1 Head Localization System Design

A sensor system with similar inputs as those used for the head position and orientation estimation by humans was developed, using a monocular camera and an IMU. For the

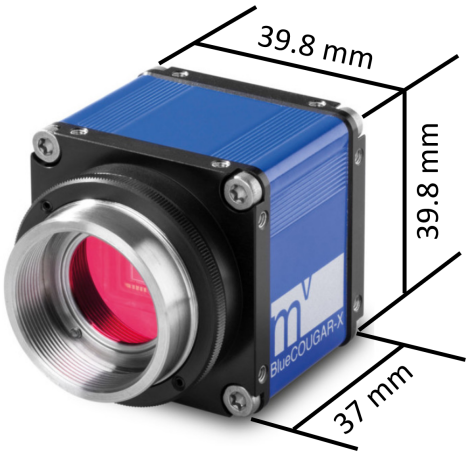


FIGURE 2.7: Camera mvBlueCOUGAR from Matrix Vision © for the visuo-inertial prototype.

TABLE 2.1: Camera parameters

Model	mvBlueCOUGAR
Frame rate	117 [Hz]
Field of View	125 [deg]
Resolution	752×480
Distortion	3%
Shutter type	Global

TABLE 2.2: IMU parameters

Model	LPMS-B
Sampling rate	400 [Hz]
Resolution	$< 0.05[deg]$
Accelerometer	3 axis $\pm(20, 40, 80, 160)[m/s^2]$, 16 bit
Gyroscope	3 axis $\pm(250, 500, 2000)[deg/s]$, 16 bit
Magnetometer	3 axis $\pm 130 \sim \pm 810[\mu T]$, 16 bit

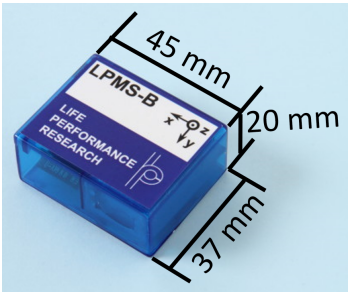


FIGURE 2.8: LPMS-B IMU from LP-research © for the visuo-inertial prototype. (From [57]).

fusion algorithm, a loosely coupled EKF based algorithm was chosen, as it may be closer to the way humans also combine visual and inertial information.

As we decided to use a monocular configuration for this work, it remains as future work to make a new sensor system using a stereo configuration, and comparing it with currently the proposed system.

2.5.2 Evaluation of the System in Simulation

Experiments using a simulated cameras and IMU were performed, using a simulator that accounts for the dynamics of the robot. To asses the performance of the estimation of the trajectories, noise was manually added to the sensor models, but the errors on the final estimation of the trajectory were decreased thanks to the use of the visuo-inertial

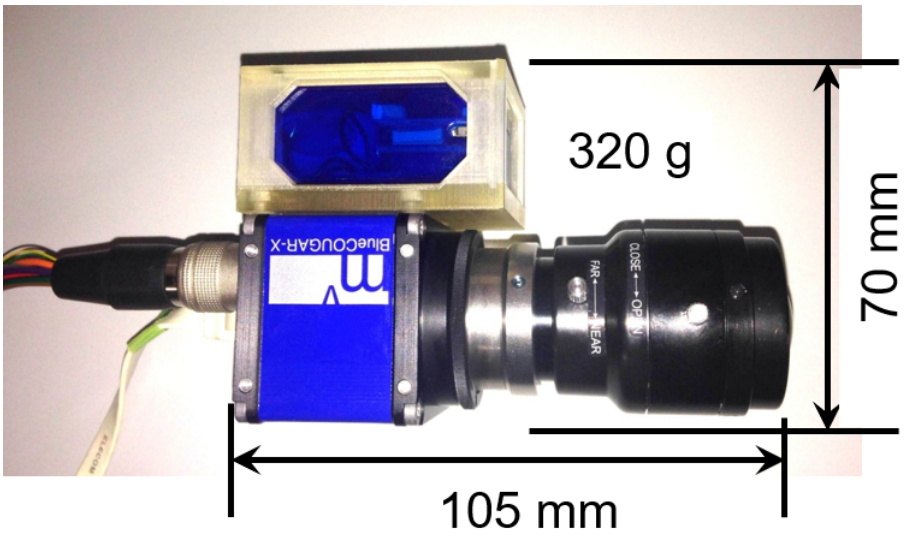


FIGURE 2.9: Visuo-inertial prototype.

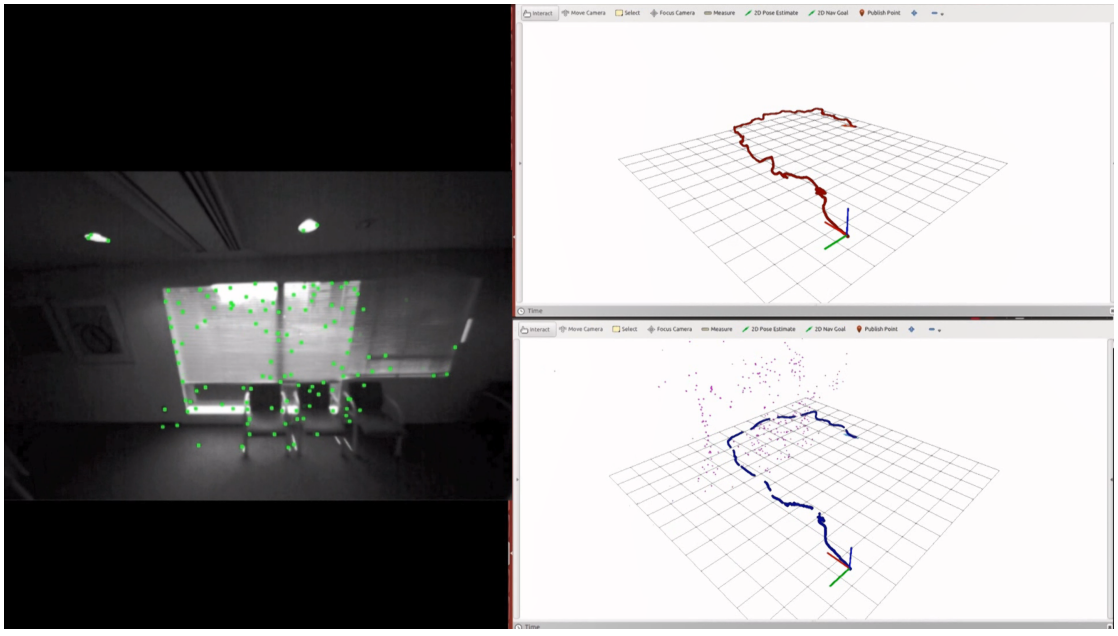


FIGURE 2.10: Snapshot of the experiment. View of the camera with found features as green points (left) and the estimated trajectory using the visuo-inertial fusion algorithm (right, top) and only VO (right, bottom).

fusion algorithm. However, as the sensors were also computational models, those results were expected, and therefore experiments with real sensors were required and performed.

2.5.3 Evaluation of the Visuo-Inertial Prototype

A visuo-inertial prototype was build and tested on a simple closed trajectory, mounting the prototype on a helmet worn by a person. The results showed that the fusion of

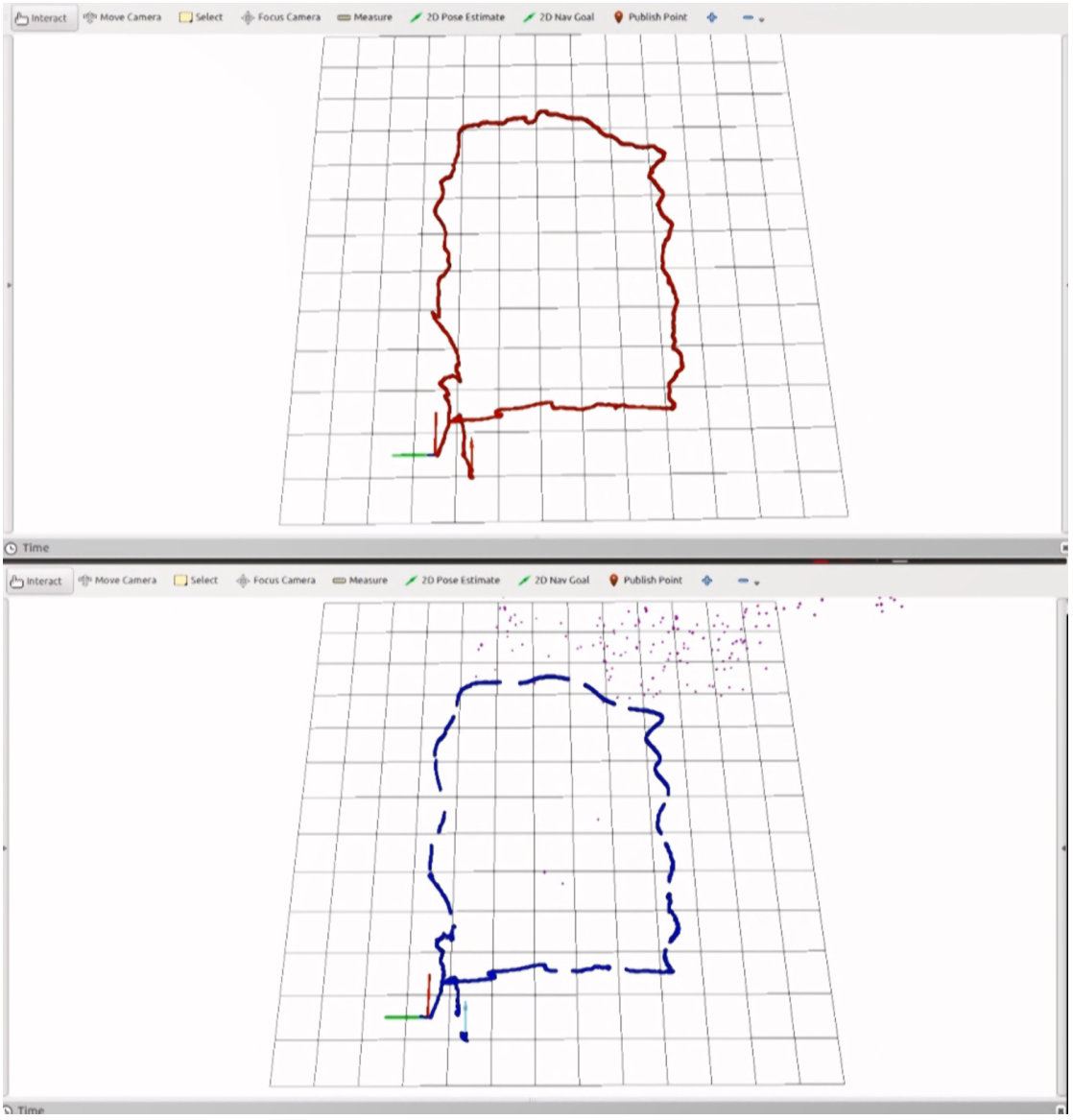


FIGURE 2.11: Estimated path of the experiment fusing visual and inertial data (up) and using only vision (down).

visual and inertial data was indeed useful for the elimination of discontinuities that appear when only using the VO algorithm, as seen in Figure 2.11. However, as the used VO algorithm was originally designed to be used for drones with down-looking cameras, it had difficulties with the walking motion, as the orientation of the camera was now forward-looking. Also, the algorithm had problems with rotational motions, so we decided to test other sparse VO algorithms for the experiments with the robot.

2.6 Summary

In this chapter we proposed a Head Localization System, comprised of a monocular camera and an IMU, for the estimation of the position and orientation of the robot's head. A loosely coupled EKF based algorithm was chosen to combine the information form the different sensors. Tests of the proposed system mounted on the robot were performed in a simulated environment, to assess the feasibility of the system, and finally a first prototype was built and tested. From the results of the prototype, we found out that the fusion algorithm was useful in order to eliminate the discontinuities that arose from estimating the robot's motion only using visual odometry. However, the used visual algorithm has problems with front-facing cameras, as well as with rotations. To solve this, we decided to explore other visual odometry algorithms, presented in the next chapter.

Chapter 3

Effects of Walking on Head Localization

3.1 Introduction

As we already mentioned in Chapter 1, the capabilities of humanoid robots must be improved for them to be able to achieve various tasks effectively and robustly in an autonomous manner for them to become more useful in our daily lives, and giving them the ability to localize themselves in the environment could be critical in this regard, and greatly help for the robot's autonomy.

One common way for the robot to self-localize is through odometry algorithms, i.e., through the estimation of the robot's change in position through the use of motion sensors, such as cameras, inertial measurement units (IMU), motor encoders, etc. These sensors can be used independently, as is the case of visual odometry (VO) algorithms, or their information can be combined to get better estimates, using algorithms as the Kalman or particle filters. The information from these odometry algorithms can then be used to send control commands to the robot depending on the application.

Once having a working self-localization estimation, it could be further improved in various ways. One approach to improve localization performance is to change the path a robot takes to a goal or the goals themselves in a way that optimizes said performance. This is called active localization, which refers to the act of partially or fully

controlling the motions of the robot to minimize the uncertainty and increase the efficiency and robustness of the estimation of its current pose [58, 59]. Humanoid robots could potentially change inter-limb coordination or gait parameters while keeping the same base trajectory, to affect the camera motion and improve the robot's localization performance.

On the other hand, as presented in Chapter 1, humans rely mainly on three inputs: visual, gravitational/inertial and proprioceptive, and in the previous chapter we introduced the indirect path model [16], which proposes the separation of the body motion estimation in two different estimations, one for the head and one for the body. In the previous chapter we developed a Head Localization System for the estimation of the position and orientation of the humanoid robot's head. Also, there are studies showing that in humans, locomotion parameters have effects on self-localization, for instance motion speed affecting our estimation of traveled distances through the use of optic flow [60]. However, current humanoid robot walking controllers and localization systems are built in ways fundamentally different from that of biological systems, and are not built purposely to achieve similar localization performance behavior (i.e. similar relationship between walking speed and localization accuracy). Moreover, previous work with hexapod robots has found inconclusive and irregular variation of SLAM performance with gait parameters [61].

From all the above, and motivated by experiments showing that humans regulate their walking speed in order to improve localization performance, in this chapter we explore the effects of walking gait on biped humanoid localization, specifically on the head localization, using the system we proposed in Chapter 2. For the regulation of walking speed on the robot, we focus on step length as it is a useful and readily applicable representation for humanoid robot locomotion planning, as in footstep planning, and because several relevant humanoid robot locomotion performance metrics as energy and slippage have been shown to depend on step length [62].

For the visual odometry algorithm, we found in the previous chapter that the used visual algorithm had problems with front-facing cameras, as well as with rotations and therefore we decided to explore other visual odometry algorithms. In this chapter we compare the performance of three different sparse VO algorithms as a function of step length: a direct, a semi-direct and an indirect algorithm.

We take a data-driven approach, i.e., we do not try to predict localization performance from simplified mechanical, control, sensor, or environment models. Instead, we directly measure localization performance of the whole system, by using ground-truth data from motion capture on many experiments while varying the robot's walking gait parameters. We focus on the following parameters:

a) Step Length

b) Bent and Stretched Knees

c) Walking Style and Symmetry

The contribution of this chapter is to answer the following questions regarding visual localization systems for biped humanoid robots:

- Does performance of such systems depend consistently and non-trivially with humanoid gait?
- What effects do different walking styles have on the performance of such systems?

In the following sections I will present the related work and the overview of the used system. Then, the performed experiments will be described, along with the presentation of the results, which will be finally discussed, comparing them to the related work and with human studies.

3.2 Related Work

Self-localization for humanoid robots has been widely researched. In the case of VO algorithms, Stasse et al. [63] proposed a real-time monocular Visual Simultaneous Localization and Mapping (VSLAM) algorithm taking into account robot kinematics from the walking pattern generator. In [64], an IMU based state estimation for a stereo based 3D SLAM is proposed, using measurements from the stereo VO and robot kinematics as updates for the Extended Kalman Filter (EKF). In [65], the authors propose an unscented Kalman filter (UKF) to estimate the ankle and hip states of a biped robot, to then use a support vector regression learning controller for bipedal walking.

Also in [66] a visuo-inertial ego-motion estimation algorithm is proposed, including the derivative of translational acceleration, also called jerk in the state vector. This was done in order to cope with the sharp accelerations generated by the ground reaction forces while walking. However, their approach did not include robot dynamics, as they tested their system with humans, rather than on a robot.

Regarding active localization using vision sensors, Davidson et al. [67] were the first to take the effects of actions into account for localization, using a stereo system attached to a mobile robot and trying to minimize the motion drift along a predefined trajectory. Also, one of the common approaches is the "Next Best View" approach, which as the name states, seeks a single additional sensor placement to reduce the localization error of the system [68]. It has been implemented as a continuous optimization method to find the whole set of future locations [69], along with a trajectory optimization using motion primitives that account for undesired motions for visual SLAM algorithms as in-place rotations [70], or together with a reinforcement learning to extract the relation between motor actions and perceptual inputs [71]. There have been works proposing different criteria to estimate the influence of the robot motions on SLAM, for example focusing on Kalman filter based approaches [72], or on the effect of the camera motion on the stability of visual localization for aerial robots [73].

An online path planning algorithm for optimal sensing using a Bayesian optimization that trades-off exploration and exploitation using a partially observed Markov decision process (POMDP) is proposed in [74].

There are also methods that use a rapidly exploring randomized tree (RRT) in conjunction with a simulated particle based SLAM algorithm to expand the tree. The simulated SLAM explicitly accounts for sensor, localization and mapping uncertainty in the planning stage [75]. In [76], they propose a path planning algorithm for planetary rovers which mainly rely on visual and wheeled odometry for localization. Their method sought to reduce the localization errors using an A* based algorithm that minimizes a cost function consisting of the distance and the trace of the covariance matrix, the latter been a way to search for feature-rich terrains.

There are works on active feature-based visual SLAM that provide real-time user feedback to minimize both map and camera pose uncertainty [77].

Also for aerial robots, methods to select paths with minimum pose uncertainty while considering the robot's dynamics have been proposed [78], as well as methods to plan paths with richer visual features [79], and more recently a method that computes the localization uncertainty optimally incorporating photometric and geometric information [59].

For legged robots, in [61] the localization accuracy of a hexapod robot was assessed, through experiments in different types of terrain and changing the robot's gait accordingly. The used sensor was an RGB-D sensor.

For biped humanoid robots, research on active visual localization has been performed attacking the problem from different perspectives, such as active localization to improve the interactions of the robot with its environment for object manipulation [80], an active vision system to estimate the location of objects while walking [81], or a task-oriented active vision system for a vision-guided bipedal walking [82]. Unfortunately, none of the above assessed the effect of the walking motion itself on the performance of the robot's localization, nor used this information to plan or modify the walking gait of the robot to obtain a better localization estimate.

From the biological point of view, humans mainly use visual, gravitational/inertial and proprioceptive cues for ego-motion estimation. Moreover, humans change the weight they put on each sensory input depending on the situation [17], [83]. Also, we know humans plan their walking gait ahead in many situations, as to keep stability in difficult situations like slippery terrains [84], but we also change our gait parameters when there are problems with the sensory inputs, as decreasing walking speed or having a more backward leaning trunk posture when visual disturbances arise [85]. Moreover, different sensory modalities perform better depending on our motions. The visual system performs better at lower frequencies than the vestibular system, but both are integrated in an optimal manner [86]. There is also evidence pointing out that modifying the walking speed has effects on our path integration abilities, making us overestimate distances when walking at slower speeds [87], as well as walking cadence affecting the performance of path integration, achieving the best performance at about 2 Hz [88].

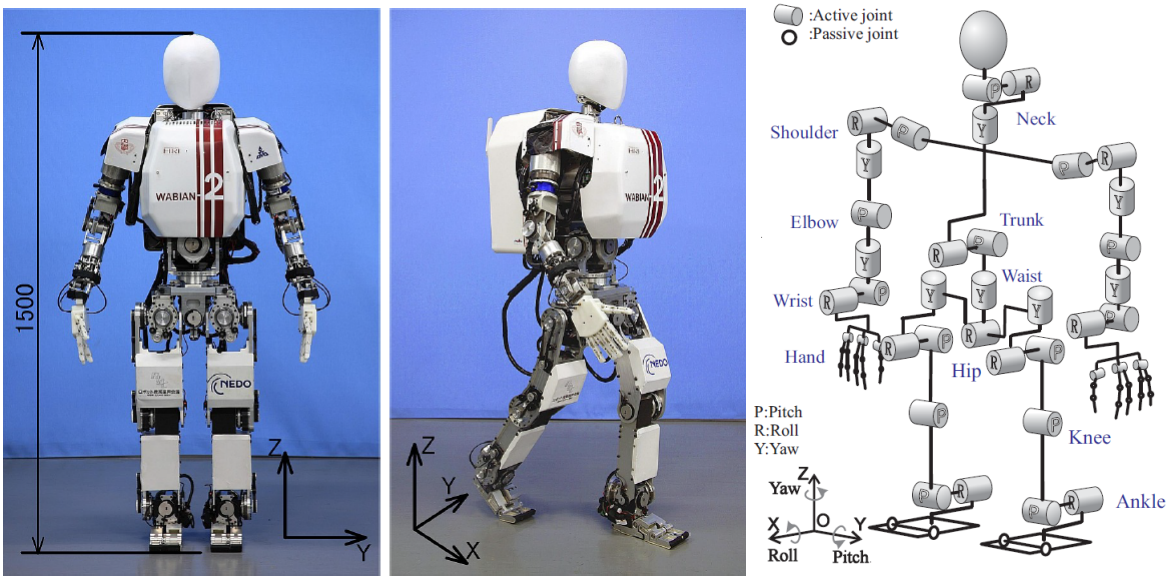


FIGURE 3.1: Robotic platform WABIAN-2R (left, center) and DoF configuration (right).

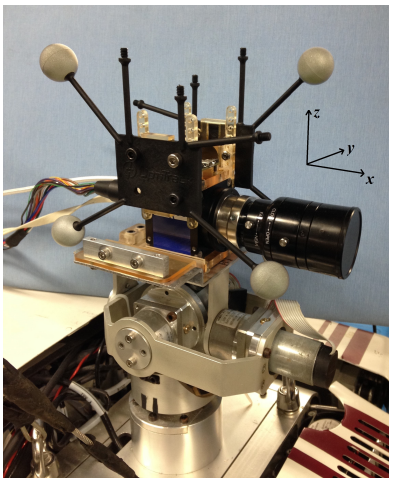
3.3 System Overview

3.3.1 Robotic Platform

In our experiments we use the humanoid robot WABIAN-2 [53], which we show in Figure 3.1. WABIAN-2 is a human-size humanoid robot, 1.5 m tall, weighting 64kg and with 33 DoFs. Joints are driven by DC-motors with high gear reduction ratios of around 200. Each motor is associated with one relative encoder and one motor driver for position control. For more details, refer to Appendix C.

3.3.2 Visuo-Inertial Setup

For the visual input, we used a Matrix Vision mvBlueCOUGAR-X (Fig. 2.7), a global shutter monocular camera, together with a low distortion wide angle lens of focal length 1.28 mm, a Field of View (FOV) of 125 deg and a distortion of 3%. The stream of images was set to 117 Hz, and the camera was mounted on the head of the robot (Fig. 3.2). For the ground truth measurements, the motion capture system OptiTrack V120:Trio at 120 fps was used, placing the photo-reflective markers on the camera to obtain the actual trajectory.



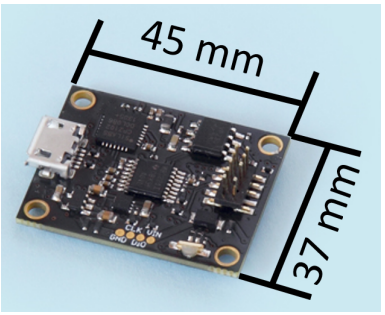


FIGURE 3.2: Close-up of the head system (head, camera, reflective markers) used for localization and ground-truth (left) and IMU[57] (right).

TABLE 3.1: Camera parameters

Model	mvBlueCOUGAR
Frame rate	117 [Hz]
Field of View	125 [deg]
Resolution	752 × 480
Distortion	3%
Shutter type	Global

TABLE 3.2: Motion capture parameters

Model	Optitrack V120:Trio
No. of cameras	3
Frame rate	117 [Hz]
Field of View	47 [deg]
Resolution	648 × 480
Shutter type	Global

TABLE 3.3: IMU parameters

Model	LPMS-CURS2
Sampling rate	400 [Hz]
Resolution	$< 0.05[deg]$
Accelerometer	3 axis $\pm(20, 40, 80, 160)[m/s^2]$, 16 bit
Gyroscope	3 axis $\pm(250, 500, 2000)[deg/s]$, 16 bit
Magnetometer	3 axis $\pm(4, 8, 12, 16)[gauss]$, 16 bit

The different reference frames and transformations used for the experiments can be seen on Fig. 3.3. We use two main reference frames, the World frame, and C_t, the frame of the camera system at time t. Also, following the notation used in [89], we define ${}^{(est)}T_{A_{t_i} \to B_{t_j}}$ as the transformation of frame B at time t_j relative to frame A at time t_i, calculated with the estimator *est*. The motion capture system tracks the camera system in the world frame, ${}^{(gt)}T_{W \to C_t}$, whereas the VO system tracks the motion of the camera system relative to its initial frame, ${}^{(vo)}T_{C_{init} \to C_t}$.

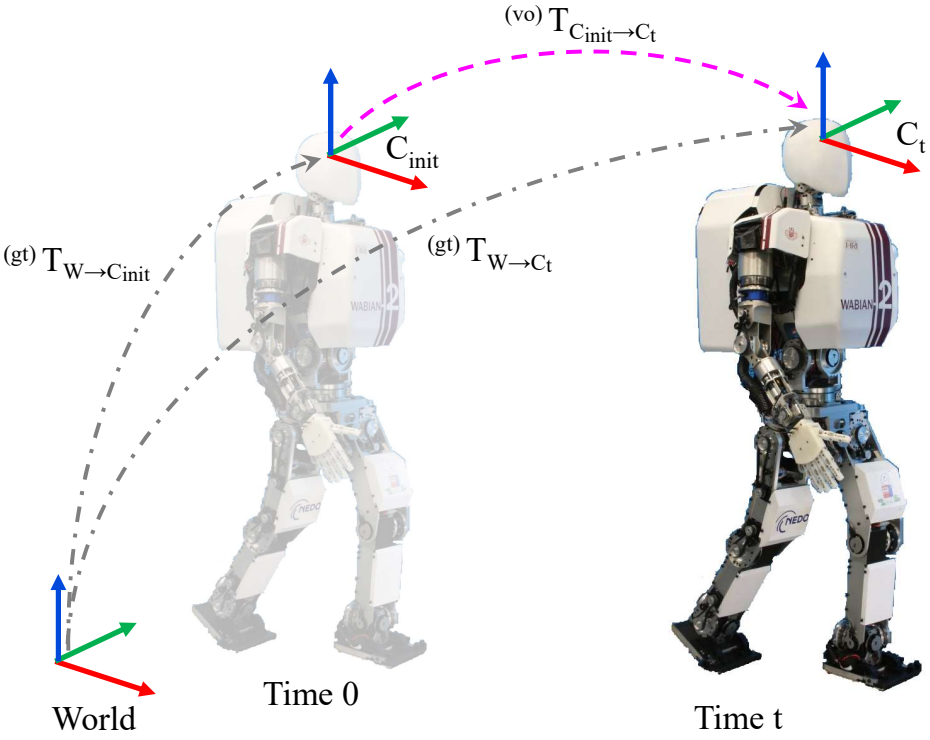


FIGURE 3.3: Used coordinate frames. C_t is the frame of the camera system at time t. The transformation of the camera frame from the world frame, obtained from the motion capture system is ${}^{(gt)}T_{W\rightarrow C_t}$. The transformation of the camera frame relative to its initial frame, calculated from the VO algorithm is ${}^{(vo)}T_{C_{init}\rightarrow C_t}$.

3.3.3 Visual Odometry Algorithms

As presented in Section 2.3, visual odometry algorithms are usually classified based on two parameters. Depending on the used information from the obtained images, the algorithms can be classified as *Direct* or *Indirect*, where *Direct* methods, as their name suggests, directly use pixel intensity information. This makes these algorithms fast, but prone to errors caused by changes in lighting conditions. On the other hand, *Indirect* methods pre-process the raw sensor data. These algorithms try to extract features such as corners, edges, or more sophisticated feature descriptors from each image, which makes them more robust to lighting changes, but computationally heavier because of the feature calculation and matching process. Another way to classify visual odometry algorithms is depending on the amount of pixels used from the input images. Based on that, the algorithms can be *Dense* or *Sparse*. *Sparse* methods use and reconstruct only a selected set of independent points, whereas *Dense* methods attempt to use and reconstruct all pixels in the 2D image domain.

As we already mentioned, we decided to test *Sparse* methods, as our focus is on localization and we do not need to reconstruct a map from the visual input. Also, as we found in the previous chapter that the used visual algorithm had problems with front-facing cameras, as well as with rotations, for this chapter we decided to explore other visual odometry algorithms. We selected and compared a *Direct* and *Sparse* method (DSO) [51], a *Semi-Direct* and *Sparse* method (SVO 2.0) [56] and an *Indirect* and *Sparse* method (ORB-SLAM2) [90].

3.3.4 Scale Extraction

Visual odometry and SLAM using only one camera suffers from the problem of scale ambiguity, i.e., with this kind of configurations it is not possible to identify the real length of translational motion only from feature correspondences [91]. To solve this, for each VO algorithm and for each experiment we calculated a scaling factor comparing the estimated traveled distance of the camera after the first step with the reference step length, assuming that ideally, given a flat floor and no slipping they should be the same:

$$\lambda_{est} = \frac{^{(ref)}d_{first\ step}}{^{(vo)}d_{first\ step}} \tag{3.1}$$

where λ_{est} is the estimated scaling factor, and $d_{first\ step}$ is the Euclidean distance between the initial position of the camera system and its position after the first step. We also chose this method as it is one of the hypothesized ways in which humans try to calculate traveled distances while walking, using substratal idiothetic cues, i.e., based on information about movement with respect to the ground or to inertial space [92].

For comparison, we also used the scale calculated using the actual traveled distance obtained from the ground truth:

$$\lambda_{real} = \frac{^{(gt)}d_{first\ step}}{^{(vo)}d_{first\ step}} \tag{3.2}$$

to examine how close was the estimated scale with respect to the real one.

3.4 Data Analysis

For the analysis of the localization performance of the different VO algorithms, we focused on the absolute trajectory error (ATE), and the relative pose error (RPE) [93]. These metrics were originally proposed for a benchmark for the evaluation of RGB-D SLAM systems. Both metrics are calculated after aligning the trajectories using the method of Horn [94], which finds the rigid-body transformation corresponding to the least-squares solution that maps the estimated trajectory onto the ground truth trajectory in closed form.

The ATE is used to asses the global consistency of the estimated trajectory, by comparing the absolute distances between the estimated and the ground truth trajectories, after both trajectories have been aligned (Fig. 5.6).

$$ATE_t = {}^{(gt)}\mathrm{T}^{-1}_{W\to C_t}\,{}^{(vo)}\mathrm{T}_{W\to C_t} \tag{3.3}$$

We then evaluated the root mean squared error over all time stamps of the translational components:

$$RMSE(ATE_t) = \left(\frac{1}{n}\sum_{i=1}^{n}\|ATE_i\|^2\right)^{\frac{1}{2}} \tag{3.4}$$

On the other hand, the RPE is used to asses the drift between the estimated and ground truth trajectories, by measuring the local accuracy of the estimated trajectory over a fixed time interval Δ. We set this time interval Δ to 10 ms, assuming that in this time interval the motion is linear (Fig. 5.7).

$$RPE_t = {}^{(gt)}\mathrm{T}^{-1}_{C_t\to C_{t+\Delta}}\,{}^{(vo)}\mathrm{T}_{C_t\to C_{t+\Delta}} \tag{3.5}$$

Similar to the ATE, we evaluate the root mean squared error over all time stamps, with $m = n - \Delta$:

$$RMSE(RPE_t) = \left(\frac{1}{m}\sum_{i=1}^{m}\|RPE_i\|^2\right)^{\frac{1}{2}} \tag{3.6}$$

3.5 Walking Speed Experiments

In this section we focus on the effects of step length, and hence walking speed, on localization performance. We generated one walking pattern for each step length, all for which the total walking distance was fixed to 1.5 m on a straight line. The used step lengths were 0.1, 0.125, 0.15, 0.175, 0.2, 0.225 and 0.25 m. Step lengths shorter than 0.1 m resulted in unstable gaits, while step lengths longer than 0.25 m were not tested since they were close to the mechanical limits of the robot. The step width was maintained constant at 0.08 m. Five runs were performed for each step length with the robot having the knees bent, i.e., maintaining a fixed height for the center of mass (CoM), and therefore for the camera. The reference walking cadence was fixed to 0.96 s/step, 0.06 seconds for double support phase and 0.9 seconds for single support phase. All patterns were executed on the robot by joint position control without any state estimation (i.e. assuming the reference trajectory of the base was executed perfectly). The motion capture and robot's joints, force, IMU and image data were stored and later analyzed.

For the visual localization, we tested three state of the art monocular visual odometry algorithms: SVO 2.0 [56], ORB-SLAM2 [90] and DSO [51], which we treated as black boxes. We fed the image stream and the intrinsic parameters of the camera, and extracted the estimated position and orientation of the camera.

We also logged acceleration and angular velocity data at 200 Hz from one IMU mounted on the camera itself, as well as force and torque data from sensors placed on both feet, also at 200 Hz. This data was processed and analyzed to look for possible differences between different walking speeds (Figs. 3.4, 3.5).

Results

We found an interesting relationship between the visual localization accuracy and the robot's step lengths used to cover the 1.5 m trajectory. The estimation from SVO 2.0 resulted to be the one with the least error, followed by ORB-SLAM2 and DSO. Interestingly, DSO seems to be affected the most by accelerations, possibly vibrations on the camera caused by the walking motions. As the step length increases, both the acceleration and the ATE of DSO increase (Fig. 3.4, 3.7). However, in Fig. 3.7 (top) the

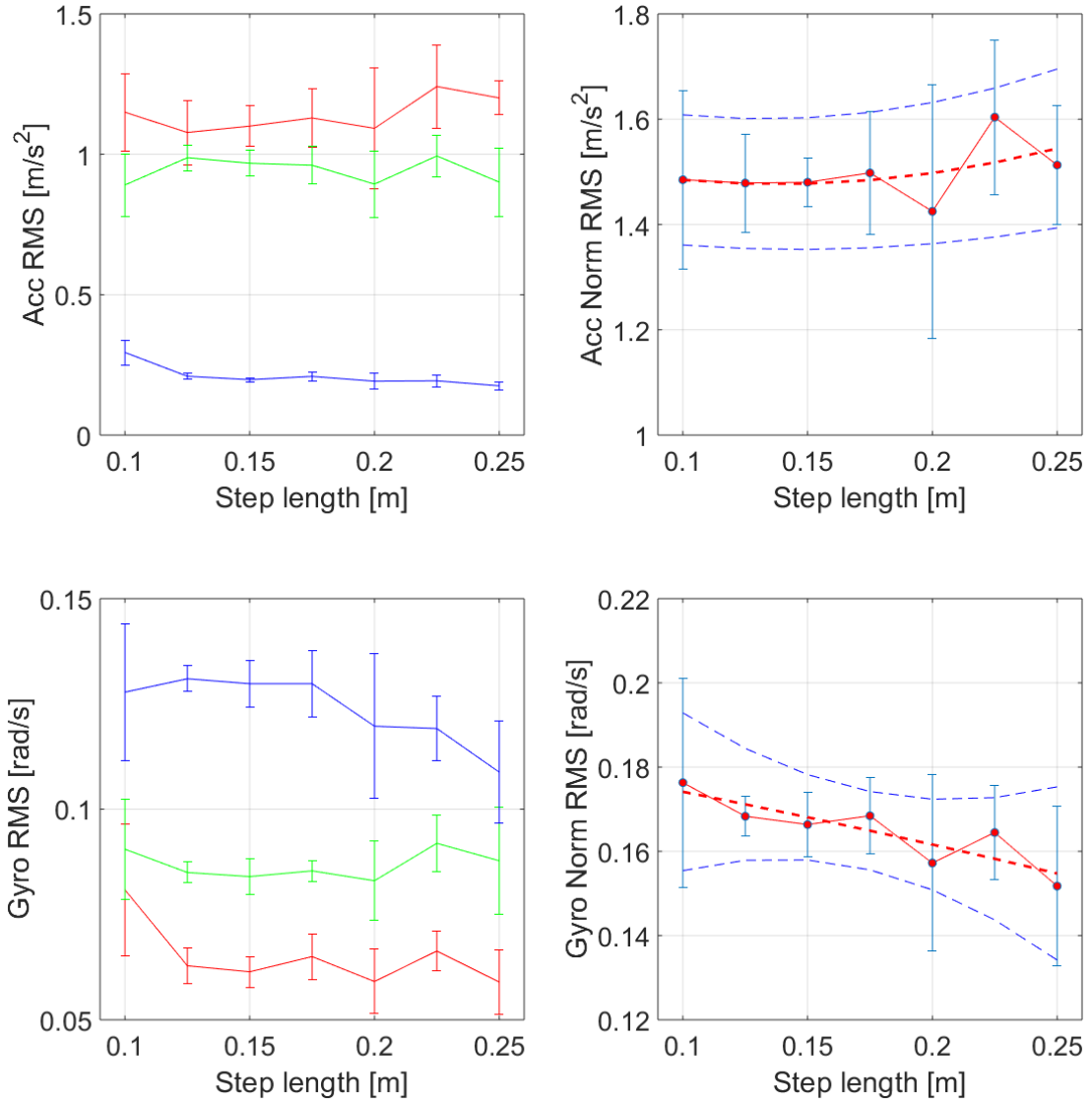


FIGURE 3.4: RMS of the data from the accelerometer and gyroscope of the IMU mounted on the camera for x (red), y (green) and z (blue) (right), and their norm, with a fitted quadratic curve (left). The used step length is on the horizontal axis.

error for a step length of 10 cm seemed like an outlier, which is also true for SVO 2.0. This coincides with the fact that with the same step length, forces in the vertical axis, as well as torques around x and y are the biggest (Fig. 3.5). This happens normally when there are early contacts of the feet with the ground or slippage, which makes the motions unstable. Also, looking at the trajectory execution error (i.e. the difference between the reference trajectory to reach the target 1.5 m away and the actual final position measured by the motion capture system), the error at a step length of 10 cm is the biggest (Fig. 3.6). Therefore, using λ_{est} for a step length of 10 cm would not be adequate, since the actual step length can be far from the actual step length. Changing the scale to λ_{real} highlighted the trend of the relationship between the performance of each VO algorithm

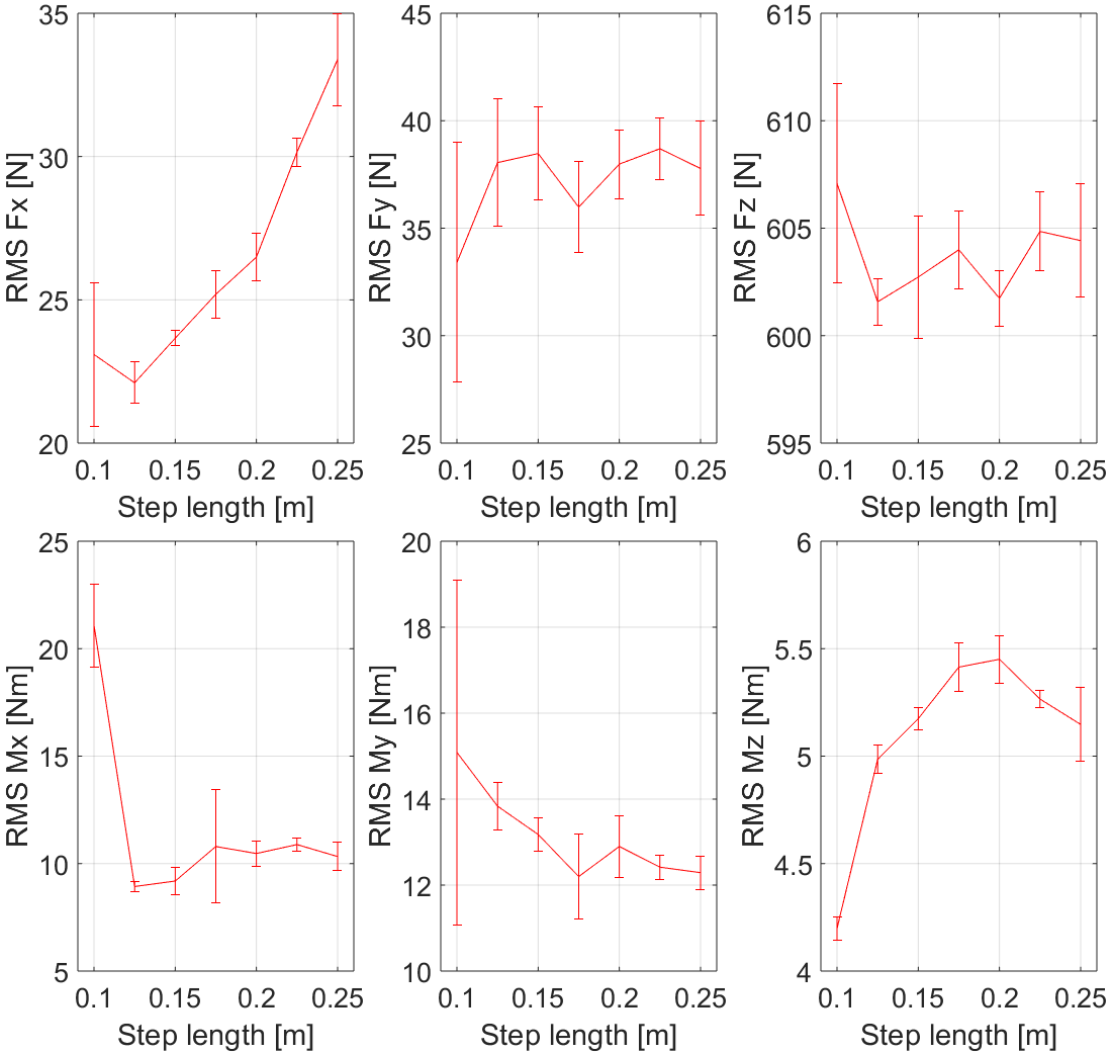


FIGURE 3.5: RMS of the data from the F/T sensors on the robot's feet. The used step length is on the horizontal axis.

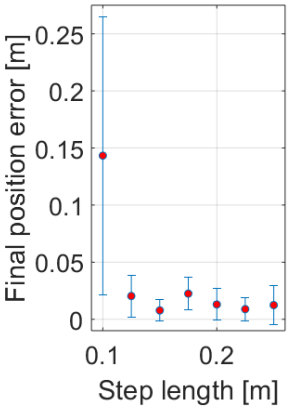


FIGURE 3.6: Trajectory tracking error, obtained with the motion capture system, w.r.t. the reference motion trajectory. The used step length is on the horizontal axis.

and step length (Fig. 3.7, bottom). From this figure, we clearly see how DSO's error increases with step length, which besides coinciding with the accelerometer data, also coincides with an increase of ground reaction forces on the feet in the x axis (Fig. 3.5).

In the case of ORB-SLAM2, the ATE decreases as the step length increases (Fig. 3.7). This could point that this algorithm, without optimizations or loop closures, is the most affected by drift. With smaller step lengths, the time to travel the reference distance is longer, which would then increase the effects of drift, whereas with longer step lengths the time is less, and so would be the drift, causing less estimation errors in the end. Finally, the trend of SVO 2.0 could be explained by the fact that, as it is a semi-direct algorithm, both effects from vibrations and drift are combining together.

To assess if the differences were significant, we also conducted a one-way ANOVA to compare the effect of step length on localization performance for each of the visual algorithms. There was significant effect for DSO at the $p < 0.05$ level [$F(6, 28) = 5.218, p = 0.001$]. Post hoc comparison using the Tukey HSD test indicated significant differences between the performances with step lengths of 10 and 20 cm ($M = -0.1037, SD = 0.0236$), 10 and 22.5 cm ($M = -0.0824, SD = 0.0236$), and 10 and 25 cm ($M = -0.1171, SD = 0.0236$) (Fig. 3.9).

A one-way ANOVA between algorithms was also performed to see if there were a significant between them (Fig. 3.10). In this analysis, we found significant effect at the $p < 0.05$ level for the step lengths of 10 [$F\,(2, 12) = 7.277, p = 0.009$], 20 [$F(2, 12) = 11.235, p = 0.002$], 22.5 [$F(2, 12) = 8.426, p = 0.005$] and 25 cm [$F(2, 12) = 19.359, p < 0.001$]. For a step length of 10 cm, the post hoc comparison using the Tukey HSD test indicated significant differences between SVO 2.0 and ORB-SLAM2 ($M = -0.0467, SD = 0.0174$), as well as between ORB-SLAM2 and DSO ($M = -0.0641, SD = 0.0174$). For a step length of 20 cm significant differences were found between SVO 2.0 and DSO ($M = -0.0727, SD = 0.0219$) and between ORB-SLAM2 and DSO ($M = -0.1003, SD = 0.0219$). For the step length of 22.5 cm significant differences were only found between ORB-SLAM2 and DSO ($M = -0.0858, SD = 0.0210$), and finally for a step length of 25 cm significant differences were found between SVO 2.0 and DSO ($M = -0.1086, SD = 0.0190$) and between ORB-SLAM2 and DSO ($M = -0.0943, SD = 0.0190$).

From the above, we observe that to minimize the effect of drift, we need to walk faster,

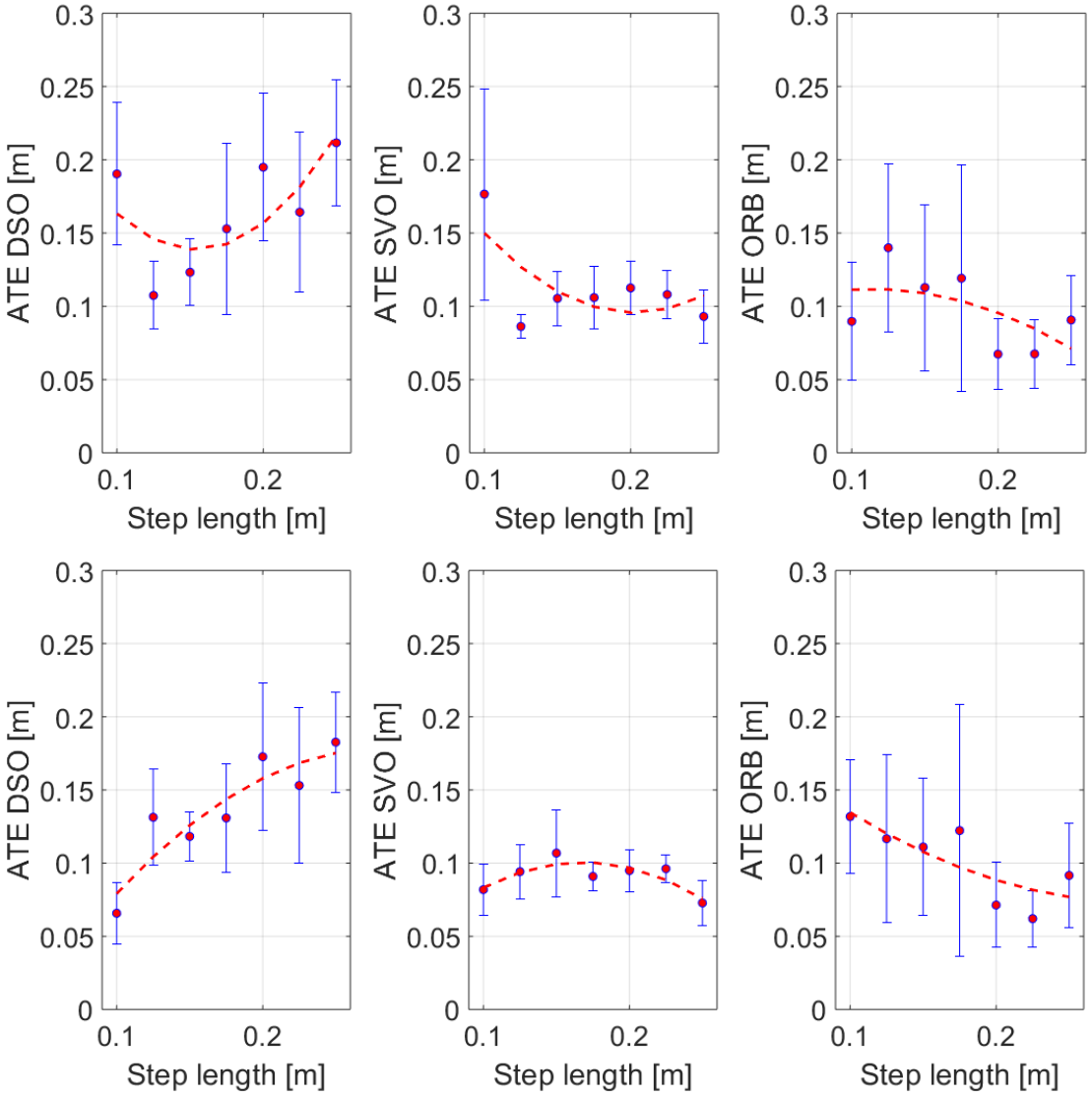


FIGURE 3.7: ATE versus step length, for DSO, SVO 2.0 and ORB-SLAM2 for λ_{est} (top) and λ_{real} (bottom). Red dots with blue vertical error bars denote the average and standard deviations for each step length, while the red dashed lines are the fitted quadratic curves for the averages. Fitted quadratic curves were calculated using the *polyfit* function of MATLAB ®.

but this in turn produces more vibrations, which negatively affect the localization performance. Therefore, we show that footstep planning could be used in an active localization system to improve the performance of VO algorithms. In this case in particular we claim that planning the footsteps so that vibrations can be minimized without too much speed decrease, could improve localization performance.

From the human point of view, there are studies showing that humans use optic flow when trying to estimate distances using vision only. Moreover, we seem to overestimate

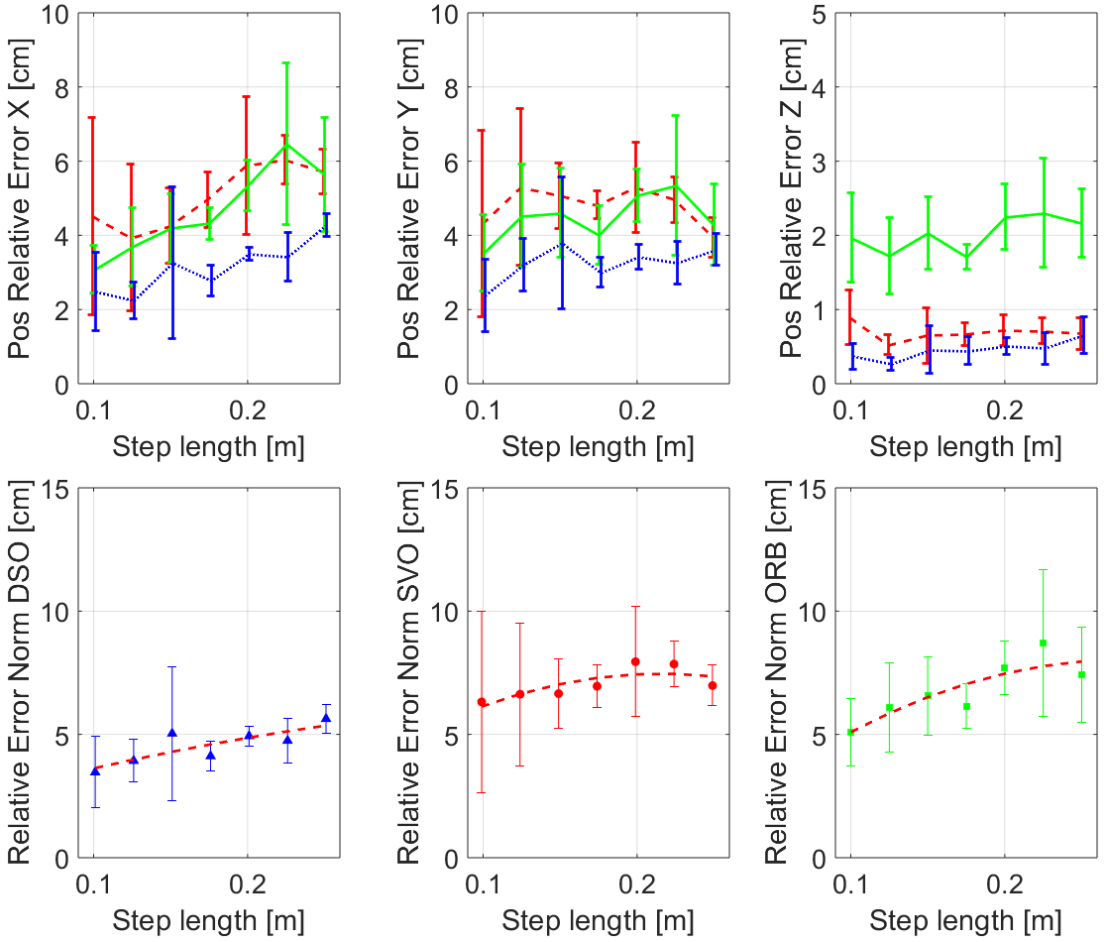


FIGURE 3.8: Relative error of the different VO algorithms w.r.t. the ground truth. SVO 2.0 (red), ORB-SLAM2 (green), DSO (blue). The used step length is on the horizontal axis.

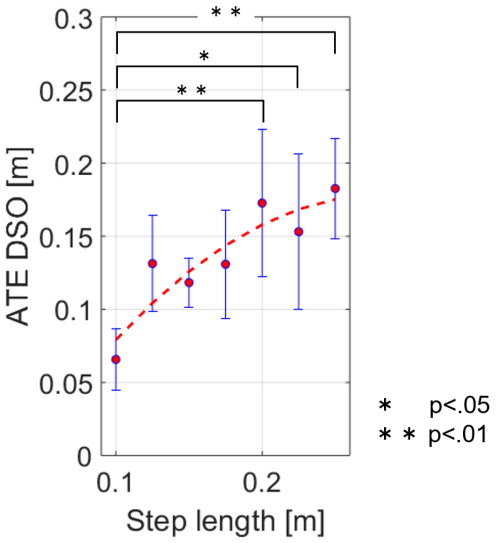


FIGURE 3.9: Statistical analysis of DSO's localization performance, measured by the the Absolute Trajectory Error .

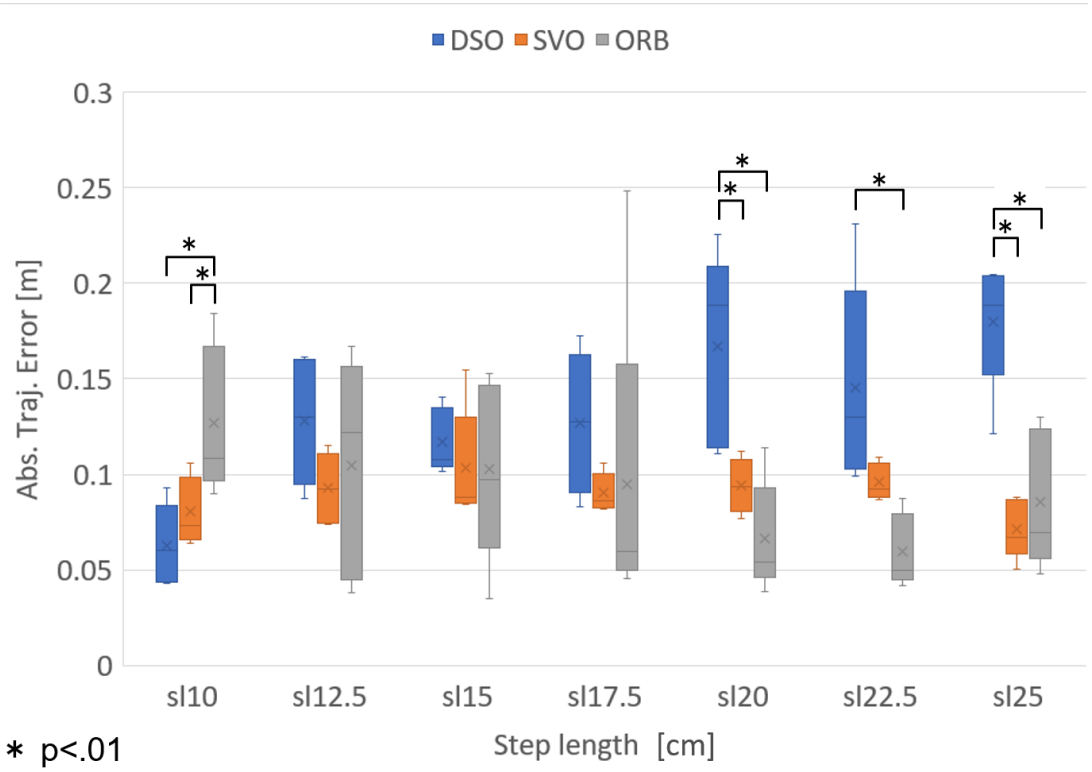


FIGURE 3.10: Statistical analysis between VO algorithms. The horizontal line inside the boxes denote the median, × denotes the mean, boxes denote the 95% confidence interval of the mean, and the whiskers denote $\pm 1 S.D.$

distances for lower speeds [60]. This coincides with the ATE of ORB-SLAM2, which increases as the step length decreases, i.e., there are more errors the slower the walking. From this, we confirmed with the robot that the behavior of distance estimation from vision for humans coincides with the behavior of a feature based VO algorithm.

As we observed a correlation between walking step length and visual localization performance, we are planning to include these localization performance curves as cost functions within a footstep planner such as to minimize localization error.

3.6 Walking with Bent and Stretched Knees Experiments

In this section we focus on the effects of walking with bent and stretched knees on localization performance. We generated one walking pattern for each step length, all for which the total walking distance was fixed to 1.5 m on a straight line. The used step lengths were 0.15, 0.2 and 0.25 m. We discarded step lengths shorter than 0.15 m as we found that they had a big trajectory execution error in the step length experiments, while

step lengths longer than 0.25 m were not tested since they were close to the mechanical limits of the robot. The step width was maintained constant at 0.08 m.

Five runs were performed for each step length for two conditions: bent knees walking and stretched knees walking. In the gait with bent knees, the robot maintained a fixed height for the CoM, and as the neck joints were fixed, the camera also maintained a fixed height; the possible variations would be produced by unevennesses on the floor. On the other hand, the gait with stretched knees introduced vertical oscillating motions to the CoM produced by the gait itself. The reference walking cadence was fixed to 0.96 s/step, 0.06 seconds for double support phase and 0.9 seconds for single support phase. All patterns were executed on the robot by joint position control without any state estimation (i.e. assuming the reference trajectory of the base was executed perfectly). The motion capture and robot's joints, force, IMU and image data were stored and later analyzed.

For the visual localization, we tested SVO 2.0 [56] and ORB-SLAM2 [90], which we treated as a black boxes. We fed the image stream and the intrinsic parameters of the camera, and extracted the estimated position and orientation of the camera.

We also logged acceleration and angular velocity data at 200 Hz from one IMU mounted on the camera itself, as well as force and torque data from sensors placed on both feet, also at 200 Hz. This data was processed and analyzed to look for possible differences between different walking speeds (Figs. 3.11, 3.12).

Results

The results from the experiments using bent and stretched knees showed no evident relationship between the visual localization accuracy and the robot's step lengths.

From the obtained results, we are planning to make experiments with more step lengths with stretched knees walking, as the current data does not show any evident difference in localization performance either walking with bent of stretched knees. However, we found that accelerations and forces in the vertical axis were higher for experiments walking with stretched knees. Therefore, we are planning to explore about the reasons behind this behavior in the future.

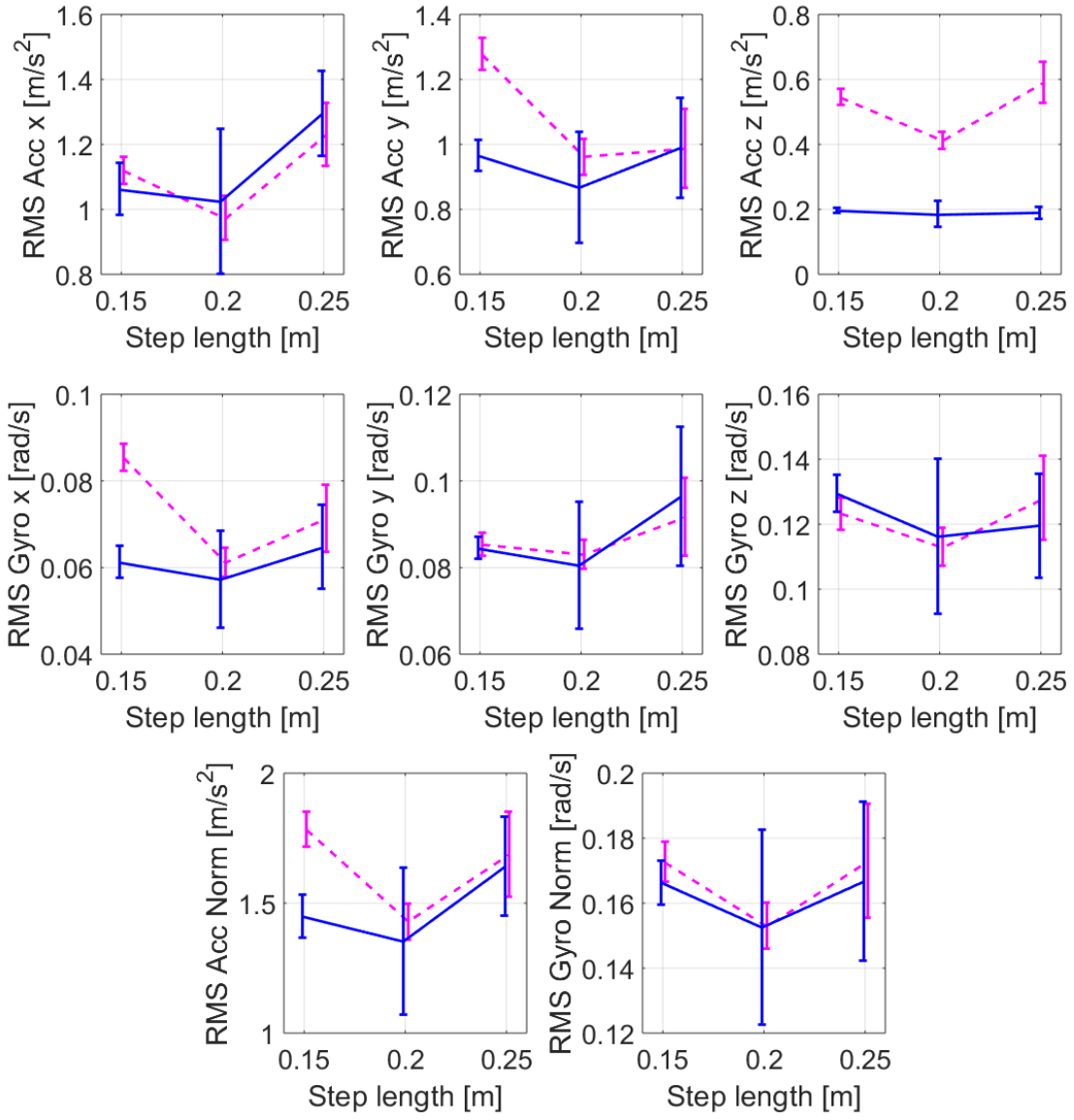


FIGURE 3.11: RMS of the data from the accelerometer and gyroscope of the IMU mounted on the camera for for bent knees (blue, solid) and stretched knees (magenta, dashed) walking experiments. The used step length is on the horizontal axis.

3.7 Walking Style and Symmetry Experiments

In this section we focus on the effects of walking style and walking symmetry on localization performance. We generated three different walking patterns, one normal walking pattern, one pattern we will call "gallop", and one we will call "slow", which will be described in the following Section. For all the patterns, the total walking distance was fixed to 1.5 m on a straight line, and the time to traverse that distance was kept inside the interval between 13.5 and 14.5 seconds. The step width was maintained constant at 0.08 m. Five runs were performed for each pattern. All patterns were executed on the

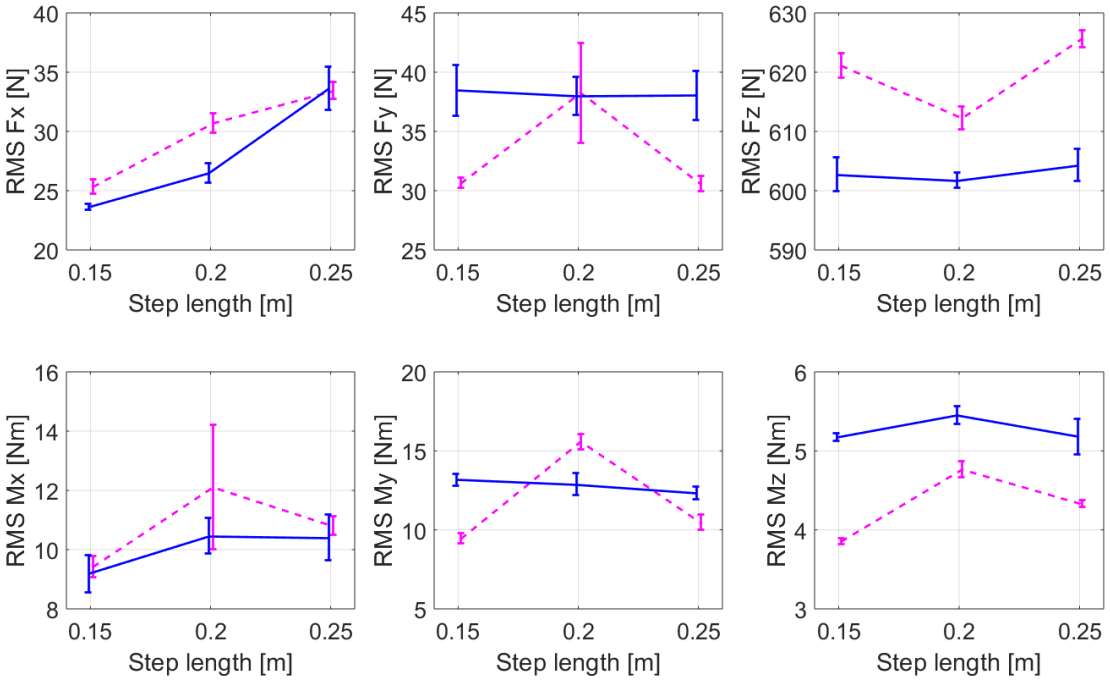


FIGURE 3.12: RMS of the data from the F/T sensors on the robot's feet for bent knees (blue, solid) and stretched knees (magenta, dashed) walking experiments. The used step length is on the horizontal axis.

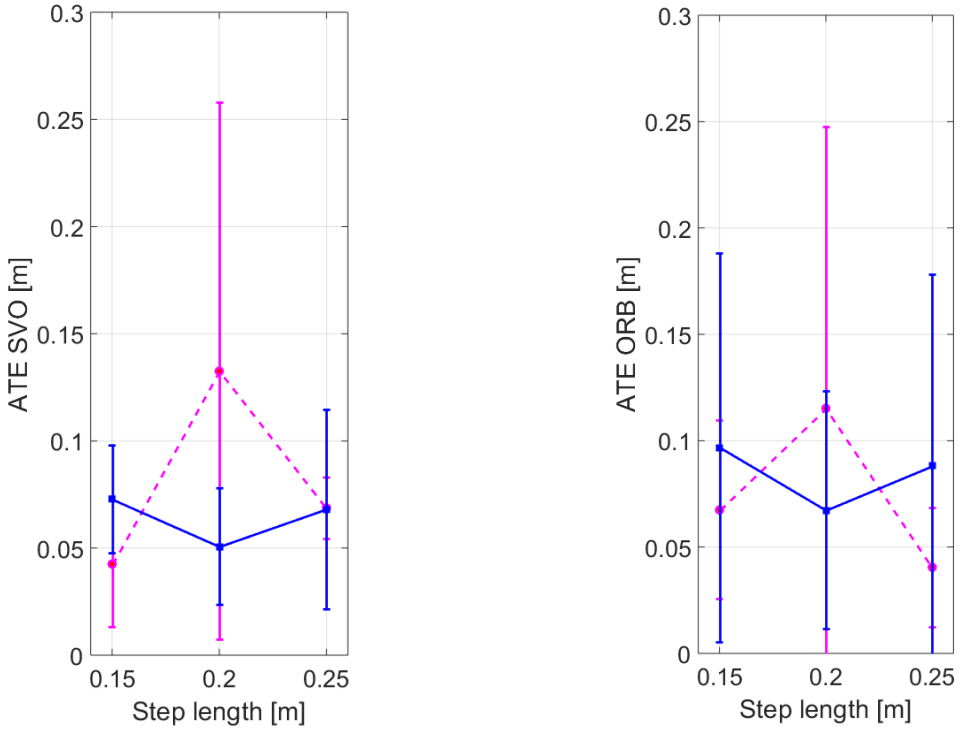


FIGURE 3.13: ATE versus step length, for DSO, SVO 2.0 and ORB-SLAM2 for λ_{est} (top) and λ_{real} (bottom), for bent knees (blue, solid) and stretched knees (magenta, dashed) walking experiments. Circles and vertical error bars denote the average and standard deviations for each step length.

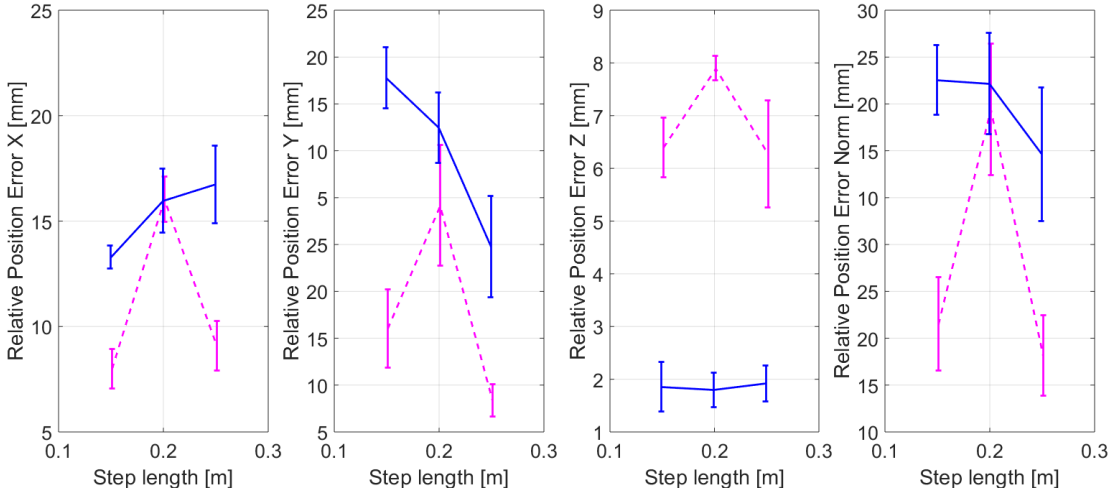


FIGURE 3.14: Relative error of SVO 2.0 for bent knees (blue, solid) and stretched knees (magenta, dashed) walking experiments. The used step length is on the horizontal axis.

robot by joint position control without any state estimation (i.e. assuming the reference trajectory of the base was executed perfectly). The motion capture and robot's joints, force, IMU and image data were stored and later analyzed.

Walking Gaits

As mentioned above, three walking patterns were tested:

- **Normal**: A walking pattern with a step length of 0.125 m and a reference walking cadence of 0.96 s/step, 0.06 seconds for double support phase and 0.9 seconds for single support phase.
- **Gallop**: A walking pattern that followed the rule 'Step forward with the right foot, then bring the left foot into alignment with the right foot, pause and repeat', as done in [95]. The step length was fixed to 0.25 m and the reference walking cadence was fixed to 0.96 s/step, 0.06 seconds for double support phase and 0.9 seconds for single support phase. (Fig. 3.15, bottom).
- **Slow**: A Normal walking pattern with a step length of 0.2 m, but a different reference walking cadence for each foot, one of 0.96 s/step (0.06 seconds for double support phase and 0.9 seconds for single support phase), and the other taking twice the time, i.e., 1.92 s/step (0.12 seconds for double support phase and 1.8 seconds for single support phase).

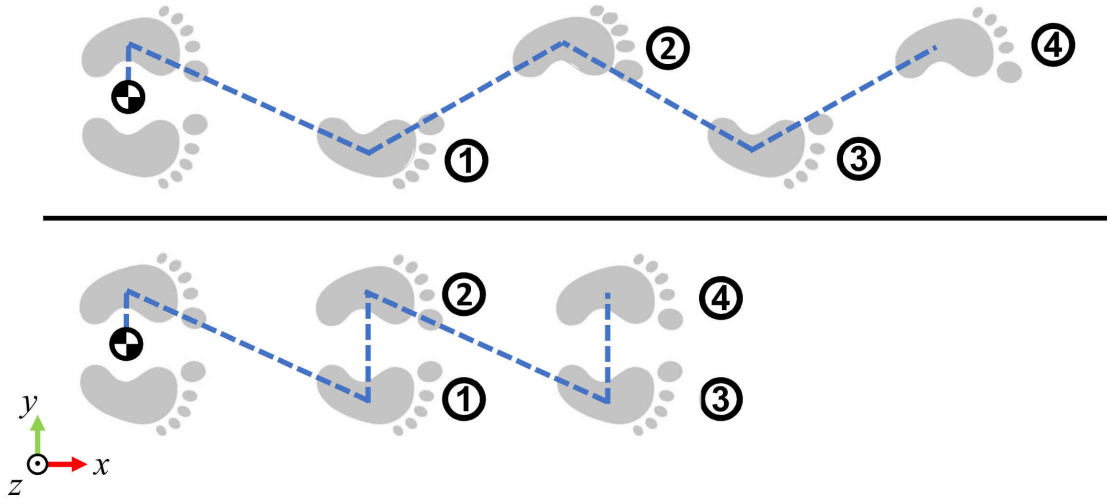


FIGURE 3.15: Stepping order for the normal and slow (top), and gallop (bottom) walking patterns and approximate zero moment point (ZMP) reference (blue dashed lines).

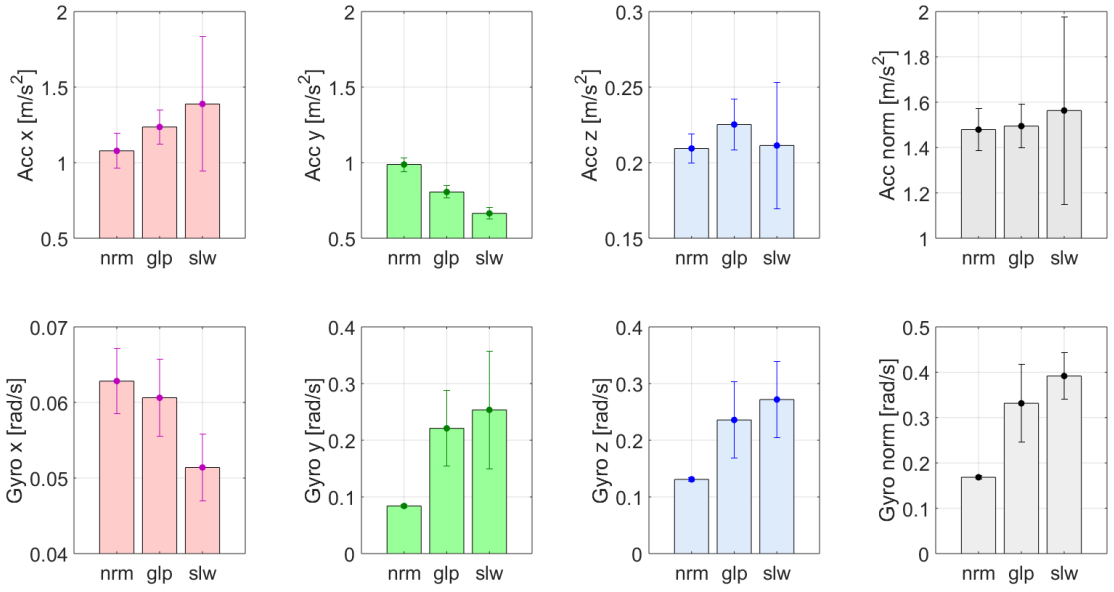


FIGURE 3.16: RMS of the data from the accelerometer and gyroscope of the IMU mounted on the camera for normal (left), gallop (center) and slow (right) in the horizontal axis. Markers with vertical error bars denote the average and standard deviations.

Results

For both visual odometry algorithms, changing the walking style from normal to gallop slightly decreased the localization error (Fig. 3.18). This could be explained by the fact that both SVO 2.0 and ORB-SLAM2 show less localization error for a step length of 0.25 m, i.e., the step length used for "gallop", than for 0.125 m, which is the one used for the normal walking gait (Fig. 3.7). Also, the moments around the y and z axes are

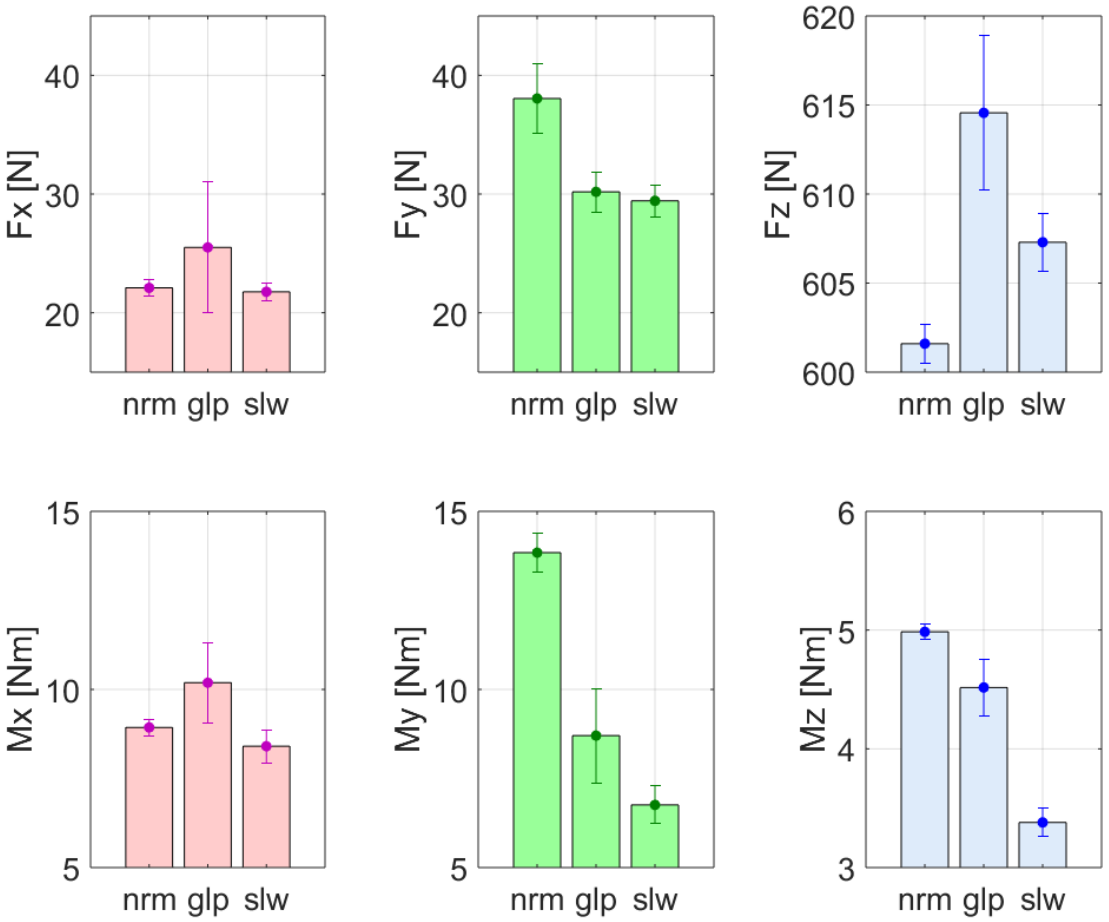


FIGURE 3.17: RMS of the data from the F/T sensors on the robot's feet for normal (left), gallop (center) and slow (right) in the horizontal axis. Markers with vertical error bars denote the average and standard deviations.

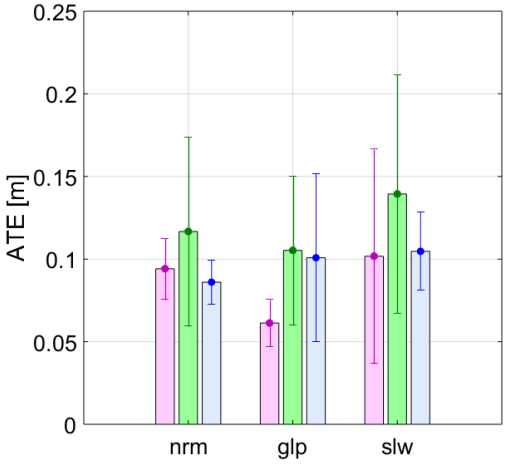


FIGURE 3.18: ATE versus walking styles for SVO 2.0 (left, magenta), ORB-SLAM2 (center, green) and kinematic odometry (right, blue). Walking styles in the horizontal axis are normal (left collection), gallop (middle collection) and slow (right collection). Markers with vertical error bars denote the average and standard deviations.

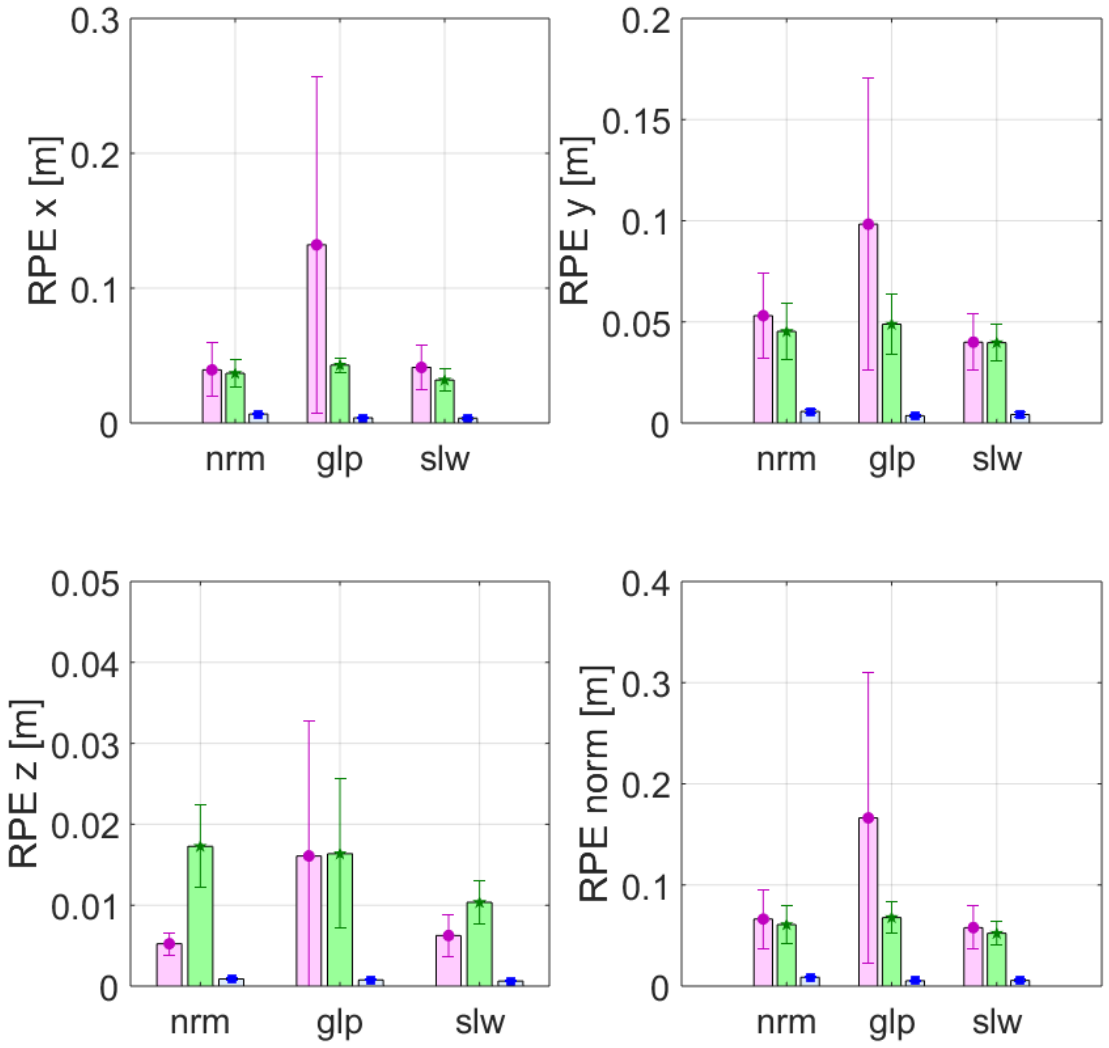


FIGURE 3.19: RPE versus walking styles for SVO 2.0 (left, magenta), ORB-SLAM2 (center, green) and kinematic odometry (right, blue). Walking styles in the horizontal axis are normal (left collection), gallop (middle collection) and slow (right collection). Markers with vertical error bars denote the average and standard deviations.

smaller for "gallop" than for "normal" (Fig. 3.17), which could be another reason for the improvement on the localization performance.

On the other hand, changing from a normal to an asymmetrical gait ("slow" gait) increased both the error as well as the variance of the visual localization. From Fig. 3.7, and given that the step length for "slow" was 0.2 m, we could expect the error for SVO 2.0 to be similar, and for ORB-SLAM2 to be smaller. However, ORB-SLAM2 is strongly affected by rotations, and in this case we can see high angular velocities for "slow" in the y and z axes (Fig. 3.16, lower row). In the case of SVO 2.0, the increase of localization errors could be caused by the high accelerations in the x axis, as well as the high variance of the accelerations on the z axis (Fig. 3.16, upper row).

It is interesting to note, however, that the moments around all the axes were the smallest for this asymmetrical gait (Fig. 3.17, lower row), but it did not seem to improve the performance of any localization algorithm.

3.8 Discussion

We will organize the discussion of this chapter according to the objectives set in the beginning and the different performed experiments.

3.8.1 Walking Speed

As already presented in Section 3.5, we found different localization performances depending on the type of the used visual odometry algorithm. For the direct algorithm, the visual localization performance got worse the faster the robot walked (Fig. 3.7, bottom left). Our hypothesis is that increased accelerations are affecting this algorithm, as we can see that accelerations also increase with step length (Fig. 3.4, top). These accelerations could be caused by the increase of vibrations due to bigger impacts of the feet with the ground as the step length increases. These vibrations in turn could be introducing motion blur to the images, which would affect the estimations from the direct VO algorithm, as it directly uses the pixel intensity information to estimate the camera's trajectory. Interestingly, ground reaction forces on the feet in the x axis also increased with step length (Fig. 3.5), which could also be another reason for the increase of estimation error for the direct VO algorithm.

On the other hand, the indirect algorithm showed a the opposed behavior to the direct algorithm, i.e., the visual localization performance got worse the slower the robot walked (Fig. 3.7, bottom right). Although the statistical analysis did not show a significant difference between the performance at different step lengths, we attribute this behavior to errors accumulating due to drift. With smaller step lengths, the time to travel the reference distance is longer, making the algorithm accumulate more errors, whereas with longer step lengths the time is less, and so would be the accumulated error.

The semi-direct algorithm showed a good performance throughout all the used step lengths, showing no evident effect of walking on its performance (Fig. 3.7, bottom

center). However, out hypothesis is that being a semi-direct algorithm, it has the advantages from both direct and indirect methods. However, that could also make this algorithm have disadvantages from both direct and indirect methods. From [56], we know that this semi-direct algorithm extract features and descriptors, i.e., acts as an indirect method, only for keyframes, and not for every frames, while the image alignment, which is performed for every frame, is performed as a direct algorithm, by minimizing the intensity differences, also called photometric error, of pixels observing the same 3-D point. Therefore, it remains as a future work to test this semi-direct VO algorithm changing the keyframe generation parameter, which would change the how direct of indirect would the algorithm behave, i.e., the more generated keyframes, the more "indirect" the algorithm, and vice-versa.

From the statistical analysis, we can propose the use of different VO algorithms for different step lengths. For a step length of 10 cm, the direct and semi-direct methods perform significantly better than the indirect method. For the step lengths of 12.5, 15 and 17.5 cm there was no significant difference, and any algorithm could be used. For the step lengths of 20 and 25 cm, the indirect and semi-direct VO algorithms performed significantly better than the direct method, and for the step length of 22.5 cm, only the indirect algorithms performed significantly better than the direct method. From these results, a possible future work is to develop a VO based localization system which changes the way of estimation, depending on the motion speed. One interesting option could be the modification of the used semi-direct method [56], in order to modify the "directness" of the algorithm depending on the motion speed, as we mentioned above.

However, it remains as future work to evaluate other visual odometry algorithms of each kind, in order to find if the relationships found in this study really depend on the type of algorithm (direct, semi-direct, indirect), of if these depend on each individual algorithm. Also, as we observed a correlation between walking step length and visual localization performance, we are planning to include these visual localization performance curves as cost functions within a footstep planner such as to minimize localization error.

3.8.2 Bent and Stretched Knees

For this experiments, we were expecting to find more errors on the stretched knees walking overall, compared to the bent knees walking. The reasoning behind this hypothesis

was simple: the walking patterns using bent knees should not have, ideally, any motion of the camera system on the vertical axis, and therefore the camera system should remain more stable, resulting in less blurring on the images and obtaining a smoother and more accurate trajectory estimation. On the other hand, walking with stretched knees necessarily introduces a vertical swaying motion, which we hypothesized that it would introduce more errors to the estimation due to making the trajectory more complex. This hypothesis was reinforced as we were not able to obtain good trajectory estimations using the direct visual odometry algorithm, which kept failing for the stretched knees experiments, and this could also be explained by the introduction of motion blur due to the vertical swaying motions of the stretched knees walking. We found some possible evidence towards the mentioned hypothesis from the analysis of the vertical acceleration and force measurements, which were higher for all step lengths when walking with stretched knees. The increased acceleration readings can be explained by the introduction of vertical motions during walking with stretched knees, and these vertical motions could also be the reason behind higher vertical force readings, caused by stronger impacting of the feet with the ground while walking. However, regardless of the aforementioned vertical acceleration and force variations, we did not find any evident difference between the localization performance of the different visual odometry algorithms. Therefore, as we already mentioned, we need to make experiments to reassess the existence of any difference on the localization performance depending on the knees behavior, and also to better understand the reason behind the higher accelerations and forces on the vertical axis for walking with stretched knees.

3.8.3 Style and Symmetry

Inspired from [95], we hypothesized that perhaps visual localization algorithms would also be affected by walking style and symmetry.

In the case of walking style, the results showed less localization error using the gallop gait with a step length of 0.25 m than using a normal gait with a step length of 0.125 m (Fig. 3.18). However, further exploration is needed to clarify if this difference is caused by the different walking styles, or by the different step lengths, as for both VO algorithms, longer step lengths tended to have less localization error (Fig. 3.7).

Regarding walking symmetry, temporal asymmetry affected negatively the visual localization performance for both the semi-direct and indirect algorithms, even though from the step length point of view it should have improved. For the semi-direct algorithm, the hypothesis is that higher accelerations in the x axis caused the increase of localization errors, whereas for the indirect algorithm, high angular velocities in the y and z axes could have been responsible for the lower performance of the visual localization with a walking gait with temporal asymmetry.

As we observed a correlation between walking style and localization performance, we are planning to include these localization performance curves as cost functions within footstep planners [96] in order to minimize localization error, just as in the case of walking speed. Moreover, as this time we focused on temporal asymmetry, we are planning to explore other kinds of asymmetries, such as posture asymmetry for future experiments.

3.9 Summary

In this chapter we explored how different walking parameters affect the performance of visual odometry algorithms, focusing on sparse methods as we are assessing the performance of localization rather than mapping. The explored parameters are step length, bent and stretched knees, and walking style and symmetry.

To asses the effects of walking speed we performed a set of experiments with a biped humanoid robot changing the step length of the walking patterns, in order to find out whether this parameter would affect the performance of VO algorithms. As we are focusing on localization and not mapping algorithms, we tested a Direct (DSO), a Semi-direct (SVO2.0) and an Indirect (ORB-SLAM2) VO algorithm. Also, we chose the step length as a proxy for walking velocity. Increasing the step length of the walking gaits showed an increase on the acceleration measurements, most likely because faster walking introduced more vibrations on the robot, which affected the performance of DSO. Also, we observed worse localization performances for ORB-SLAM2 the shorter the step lengths, i.e., the slower the walking, and as it took more time to get to the goal, the effect of drift on the localization estimates was also increased.

Then, to asses the effects of walking with bent and stretched knees we performed a set of experiments with bent knees and stretched knees. We assessed whether these

parameters would affect the performance of two VO algorithms, a Semi-direct (SVO2.0) and an Indirect (ORB-SLAM2) algorithm. The Direct (DSO) algorithm was not robust enough, so we did not include it in this study. However, the obtained results did not show any evident difference in localization performance either walking with bent of stretched knees. Therefore we should make experiments with more step lengths with stretched knees walking to be able to analyze with more details the behavior of such walking styles.

Finally, to asses the effects of walking symmetry and style we performed a set of experiments for different walking styles and gait symmetry conditions, in order to find out whether these parameters would affect the performance of visual and/or kinematic localization. We tested a Semi-direct (SVO2.0) and an indirect (ORB-SLAM2) VO algorithms.

Using a gallop gait decreased the localization error for visual localization, which the data shows to be related to a decrease in the moments around y and z, caused either by the walking style itself, or because of the change in step length.

Eliminating the temporal symmetry of the walking gait increased the error of the visual localization, as well as its variance, even when from the step length point of view the error should have either remained or improved. For ORB-SLAM2 rotations could have affected the performance, whereas for SVO 2.0 accelerations, most likely produced by vibrations during walking, affected its performance.

Chapter 4

Body Localization

4.1 Introduction

As we already mentioned, we are designing a localization system inpired by the the indirect path model [16], which proposes the separation of the body motion estimation in two different estimations, one for the position and orientation of the head, and one for the position and orientation of the body. In this chapter we will now present the development of the Body Localization System, the system for the estimation of the position and orientation of the humanoid robot's body.

We will present the possible sensors to use, in order to emulate the same inputs that humans use for this task, as well as the used method to combine the information from these sensors to obtain an estimation of the position and orientation of the system. With that in mind, the objectives for this chapter are as follows:

a) Design a sensor system with similar inputs as those used for the body position and orientation estimation by humans

   • Choose the adequate sensors for the system, to obtain similar information as that from human senses

   • Decide the fusion algorithm to be used to combine the information from the chosen sensors

b) Test the validity of the proposed system

In order to achieve the objectives, the proposed system will be presented in the following sections, first introducing the chosen sensors as well as the selected fusion algorithm. Then, simulation tests and results of the proposed system and are shown. Finally validation experiments with a real humanoid robot are presented, to evaluate the performance of the proposed system in the real world through simple experiments.

4.2 Sensors

According to the indirect path model [16], the required information to estimate the position and orientation of the body is an estimate of the orientation of the body in space, given by the somatic graviceptors, and proprioceptive information from the body.

To emulate the different sensory inputs, we chose sensors that would provide us with similar data. Therefore, for the somatic graviceptive inputs we chose IMUs and for the proprioceptive inputs, we focused on sensing joint position and motion, using optical encoders, and the distribution of the bodyweight on the feet, using six axis force/torque sensors.

4.3 Kinematic and Inertial Fusion

For the localization of the body, we decided to use a probabilistic fusion algorithm to combine kinematic and inertial information. For this, we used the method first developed by Bry et al. [97] and extended for biped robots by Fallon et al. [98], which will be briefly described in the following paragraphs. The method is also based on an EKF, where the process model is based on a discrete time, nonlinear discrete transition function

The state vector is defined as:

$$x^T = \begin{bmatrix} {\omega_b}^T & {v_b}^T & {R_W^b}^T & {p_W^b}^T \end{bmatrix} \tag{4.1}$$

where ω_b is the angular velocity in body coordinates, v_b is the linear velocity in body coordinates, and R_W^b and p_W^b are the rigid body orientation rotation matrix and the translation vector from the origin in the world coordinate frame to the origin of the body frame, both expressed in the world frame.

Measurements from an IMU are used for the propagation of the states, following the equations below:

$$\dot{v}_b = -\omega_b \times v_b + {R^b_W}^T g + g\, a_m \tag{4.2}$$

$$\dot{R}^b_W = R^b_W \lfloor \omega_m \times \rfloor \tag{4.3}$$

$$\dot{p}^b_W = R^b_W v_b \tag{4.4}$$

where a_m and ω_m are the measured acceleration and angular velocity from the IMU, g is the gravity vector in the world frame, and $\lfloor v \times \rfloor$ is skew-symmetric matrix of vector v [48], defined as:

$$\lfloor v \times \rfloor = \begin{bmatrix} 0 & -v_z & v_y \\ v_z & 0 & -x \\ -v_y & v_x & 0 \end{bmatrix} \tag{4.5}$$

The angular velocity state is neglected assuming that the IMU provides accurate measurements.

The measurement update is performed using the kinematic integration from the legs, assuming a non-slipping and stationary contact of the foot with the ground during stance phase. This assumption allows to use forward kinematics to infer the instantaneous velocity and position measurements of the robot's base link. A gait transition detector is used to detect the current stage of locomotion, deciding which foot is in contact with the ground. A Schmitt trigger is used to classify the contact forces of the robot from its F/T sensors placed on the feet, and a state machine is decides which foot is in contact with the ground. This foot is then the one used as the anchor point from which the forward kinematics will be calculated.

For the measurement update, the difference between consecutive position estimates of the base link of the robot is used. The reason why the base link position is not directly used is because of the inconsistencies in joint sensing, and because the feet don't maintain perfectly static contacts with the ground. Therefore, the measurement model for the kinematics is defined as:

$$z_{kin} = \hat{v}_b = \frac{{p^b_W}^{(t)} - {p^b_W}^{(t-1)}}{\tau_s} \tag{4.6}$$

where τ_s is the integration period and p_W^b both for the current and past times are obtained from the robot's kinematics, extracting it from the transformation between the base link and the world frame T_W^b as:

$$T_W^b = \begin{bmatrix} R_W^b & p_W^b \\ 0_{1\times 3} & 1 \end{bmatrix} = \left(T_p^f\right)^{-1} T_W^f \tag{4.7}$$

where T_p^f is the transformation of the foot frame relative to the base link frame, produced by forward kinematics. T_W^f is the transformation of the foot frame relative to the world frame, defined as:

$$T_W^f = \begin{bmatrix} R_W^f & p_W^f \\ 0_{1\times 3} & 1 \end{bmatrix} \tag{4.8}$$

where p_W^f is an input from the previous time step, and R_W^f can be obtained from the rotation of the foot relative to the base link, obtained from the forward kinematics, and the current rotation of the body relative to the world, assuming that the base link orientation estimate has zero covariance:

$$R_W^f = R_b^f R_W^b \tag{4.9}$$

Additionally, in [98] they also include a measurement update using information from a LIDAR, which we are not including here as we are not using that sensor.

4.4 Robotic Platform

For the preliminary experiments, we used WALK-MAN, a biped humanoid platform developed to operate in realistic unstructured environments. We will present a brief description of the robot in the following paragraphs, but for a more detailed description please refer to [99].

WALK-MAN (Fig. 4.1) is a humanoid robot developed with three objectives:

- Powerful manipulation, with high power-to-weight ratio and reduced inertia at the legs to maximize dynamic performance.

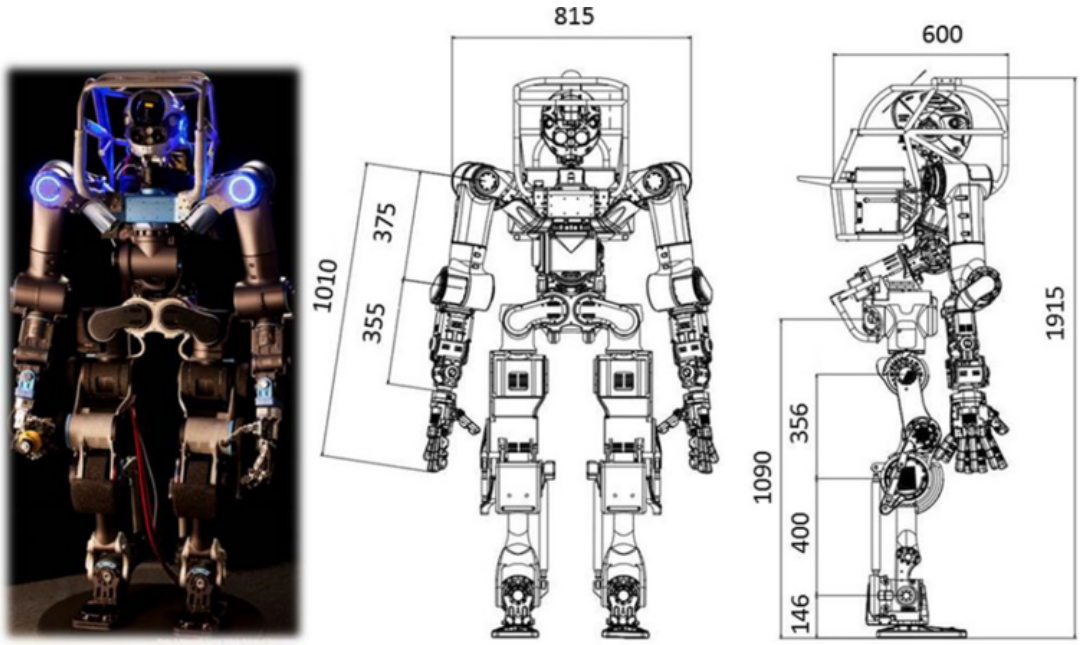


FIGURE 4.1: Robotic platform WALK-MAN, all measurements in mm (taken from [99]).

- Robust balanced locomotion, along with large joint range of motion to achieve humanlike movement.
- Enhanced physical sturdiness

It is 1.915 m high and weights 132 kg, of which 14 kg are from the power pack and 7 kg are from the protection structure. Regarding the DoFs, the upper body has 17 DoF without counting the hands, each arm having 7 DoF, with 3 DoF on the shoulder, 1 DoF on the elbow, 1 DoF for the forearm rotation an 2 DoF for the wrists. Each arm has a 19 DoF hand as end-effector. A 3 DoF waist connects the upper body with the lower body, which has 12 DoF. Each leg has 6 DoF, 3 DoF at the hip, 1 DoF on the knee and 2 DoF on the ankle. In order to reduce the leg inertia, the ankle pitch actuator was relocated close to the knee joint, implementing a 4-bar transmission mechanism to transmit the power to the ankle joint.

One of the key technologies used in the robot is the series elastic high-end actuator unit explicitly developed for the WALK-MAN project. These units consist of a frameless brushless DC motor, a harmonic drive, and a torsion bar a a flexible element, that connects the output of the harmonic drive to the output flange of the actuator. Each unit is equipped with sensors to measure joint position, torque and temperature. For the joint position, two 19-bit absolute high-resolution sensors are used, one mounted at

the output of the harmonic drive and before the elastic element, and one mounted on the link side after the elastic element. This configuration also allows the measurement of torque by monitoring the relative deflection of the torsion bar spring.

The perception system of WALK-MAN comprises several sensors to perceive both the internal states as well as the external environment.

Absolute joint position Two absolute magnetic encoders, one placed immediately after the reduction mechanism, and the other placed after the series elastic bar.

Joint torque sensing Estimated from the relative deflection of the torsion bar spring, obtained from the difference between the data from the magnetic encoders.

Force/Torque sensing Customized 6 axis load cells mounted on the ankles and wrists of the robot.

Environment sensing Multisense M7 sensor mounted on the head, which comprises a stereo vision system, a LIDAR and an IMU. Also the robot has a microphone array system around the ears to monitor sound.

Inertial and gravity sensing Additionally to the IMU inside the head system, there is a second IMU mounted on the pelvis area. These IMUs are used for the locomotion and balancing controllers.

4.5 Experiments

To test the validity of the chosen algorithm, we first performed test with logged data from past experiments with the robot, and then performed new walking experiments with ground truth measurements to assess the precision of the algorithms estimation. These will be detailed in the following sections.

4.5.1 Simulation Experiments

First, we assessed how each sensor contributed to the estimation of the position and orientation of the robot.

With only the IMU, a double integration of the acceleration is necessary to obtain the position, and as the accelerometer measurements drift over time, the integration accumulates these errors. This results in an ever drifting robot, even without any motion. This can be seen on figures 4.2 and 4.3, where as the robot drifts backwards, it seems as if it was shrinking.

With only the kinematic information, all the joints move, but as there is no information about contacts with the environment, the robot moves as if it was suspended in the air. This can be seen on figures 4.4 and 4.5, where the position of the robot w.r.t. the initial frame is constant, i.e., the robot does not move forward.

Combining kinematic information with information about the contacts with the environment, obtained from the force/torque sensors, the robot is now able to perceive it is walking, but as it has no IMU, it has no information about the inclination of the ground. For the used data, this resulted in the robot sinking into the ground. This can be seen on figures 4.6.

Finally, having the joints, force/torque and IMU data, the robot is finally able to have a better estimation of its motion. This can be seen on figures 4.7 and 4.8. The difference of the estimation between using and not usig an IMU can be seen in figure 4.9, where it can be seen that without the IMU data, the robot has no information regarding the inclination of the ground and the robot itself, and therefore in case the floor is inclined, the robot will estimate itself sinking or floating into the air.

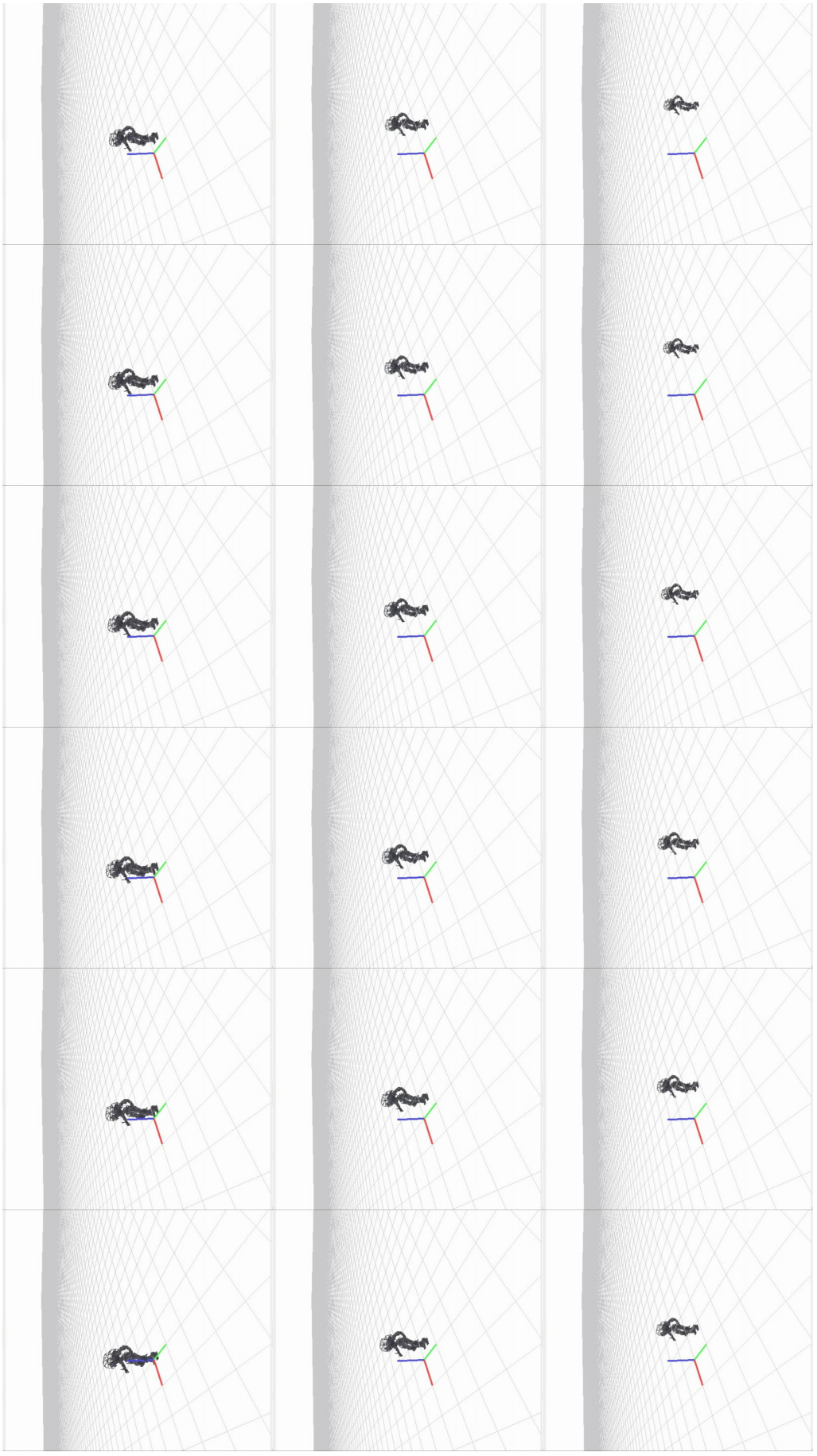


FIGURE 4.2: WALK-MAN simulation experiment: IMU only (1/2).

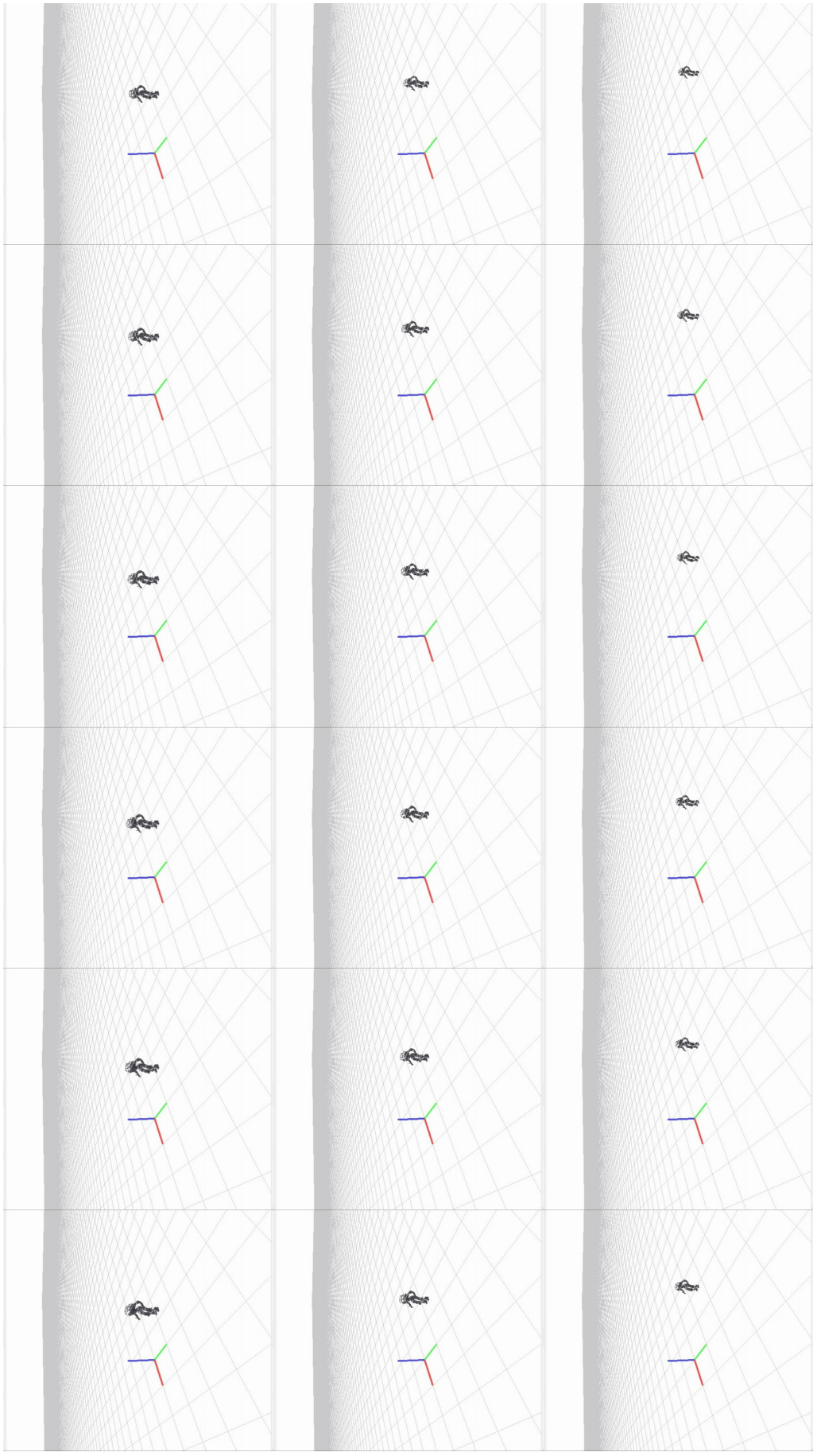


FIGURE 4.3: WALK-MAN simulation experiment: IMU only (2/2).

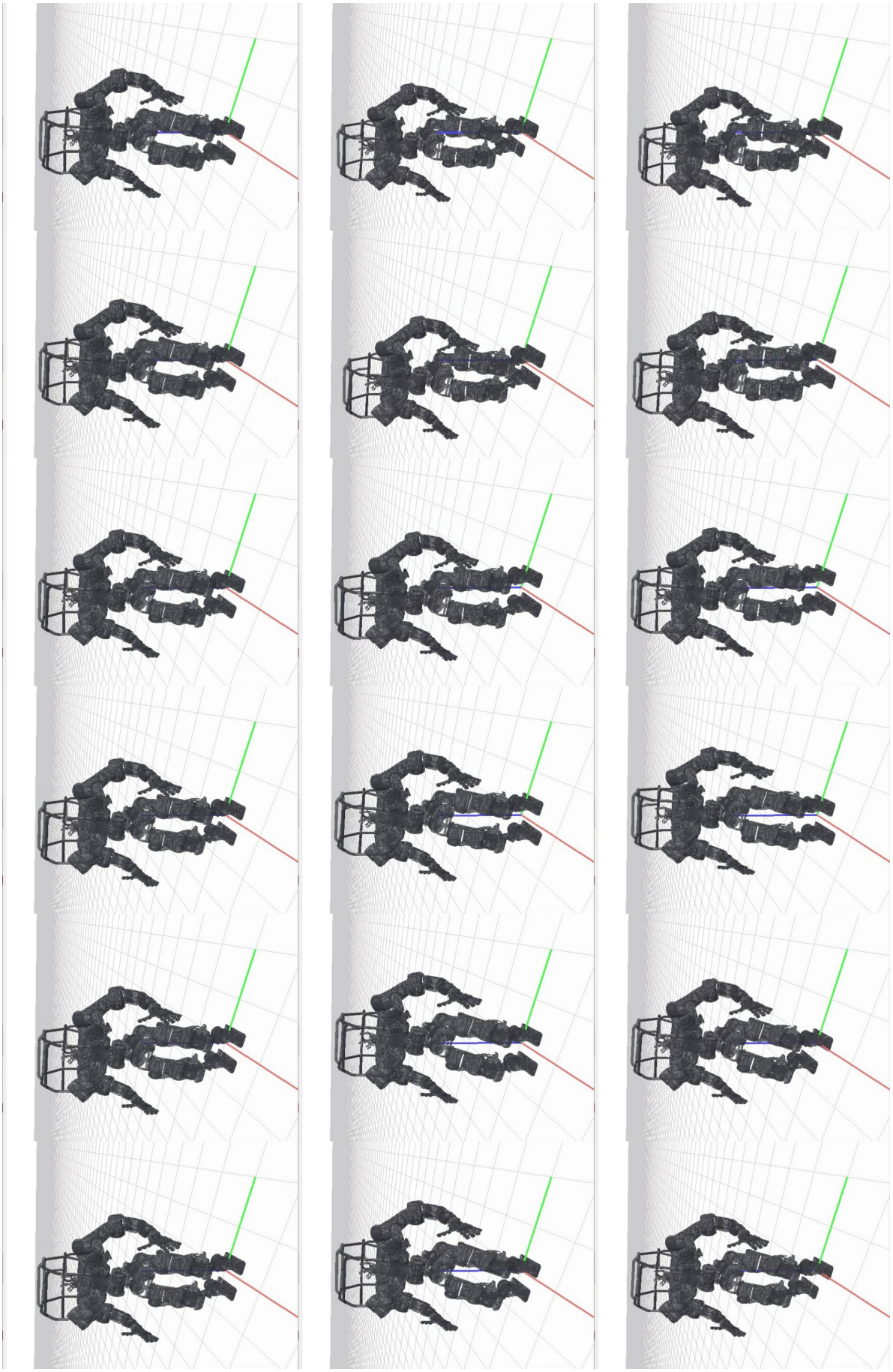


FIGURE 4.4: WALK-MAN simulation experiment: kinematics only (1/2).

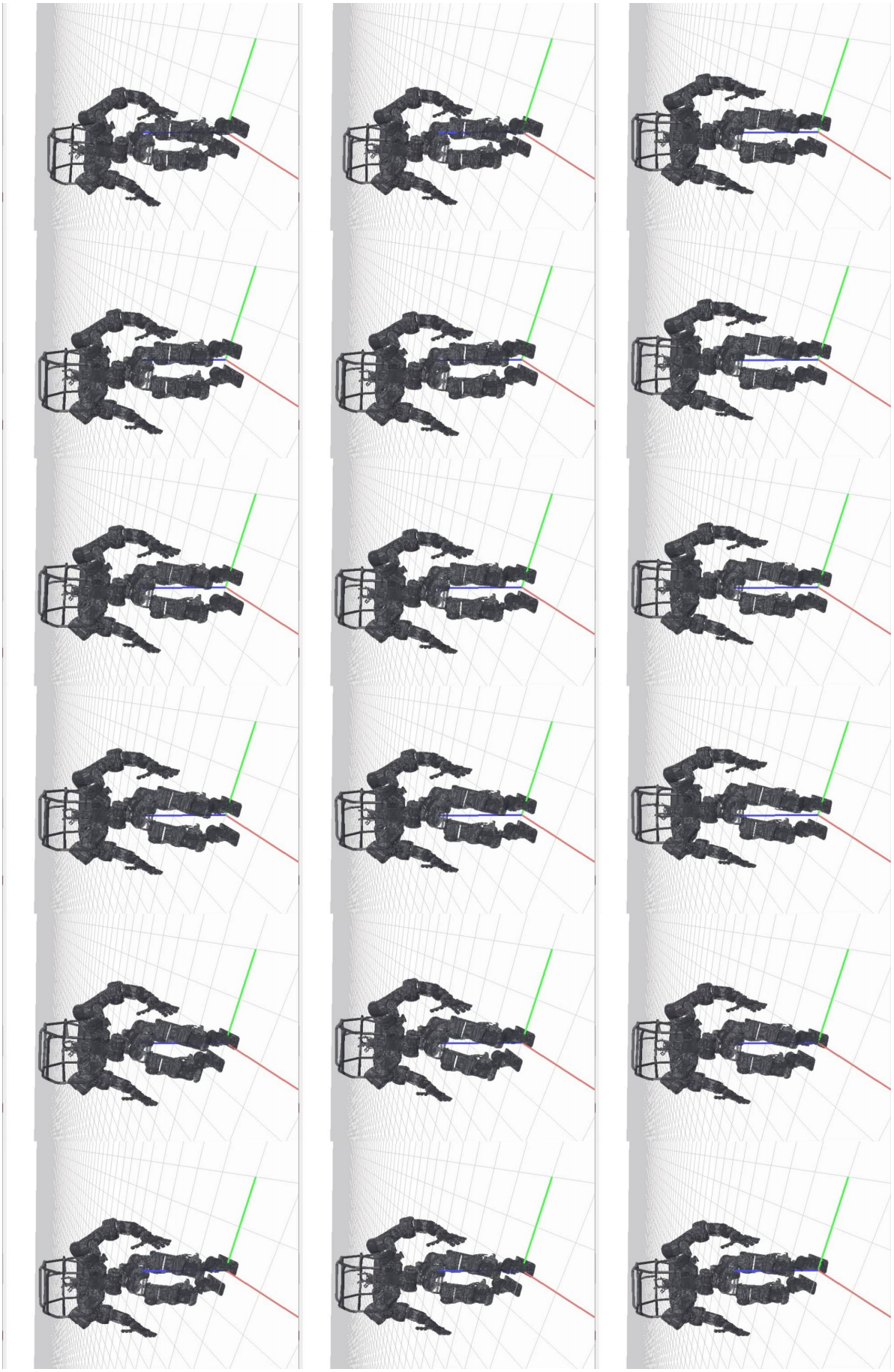


FIGURE 4.5: WALK-MAN simulation experiment: kinematics only (2/2).

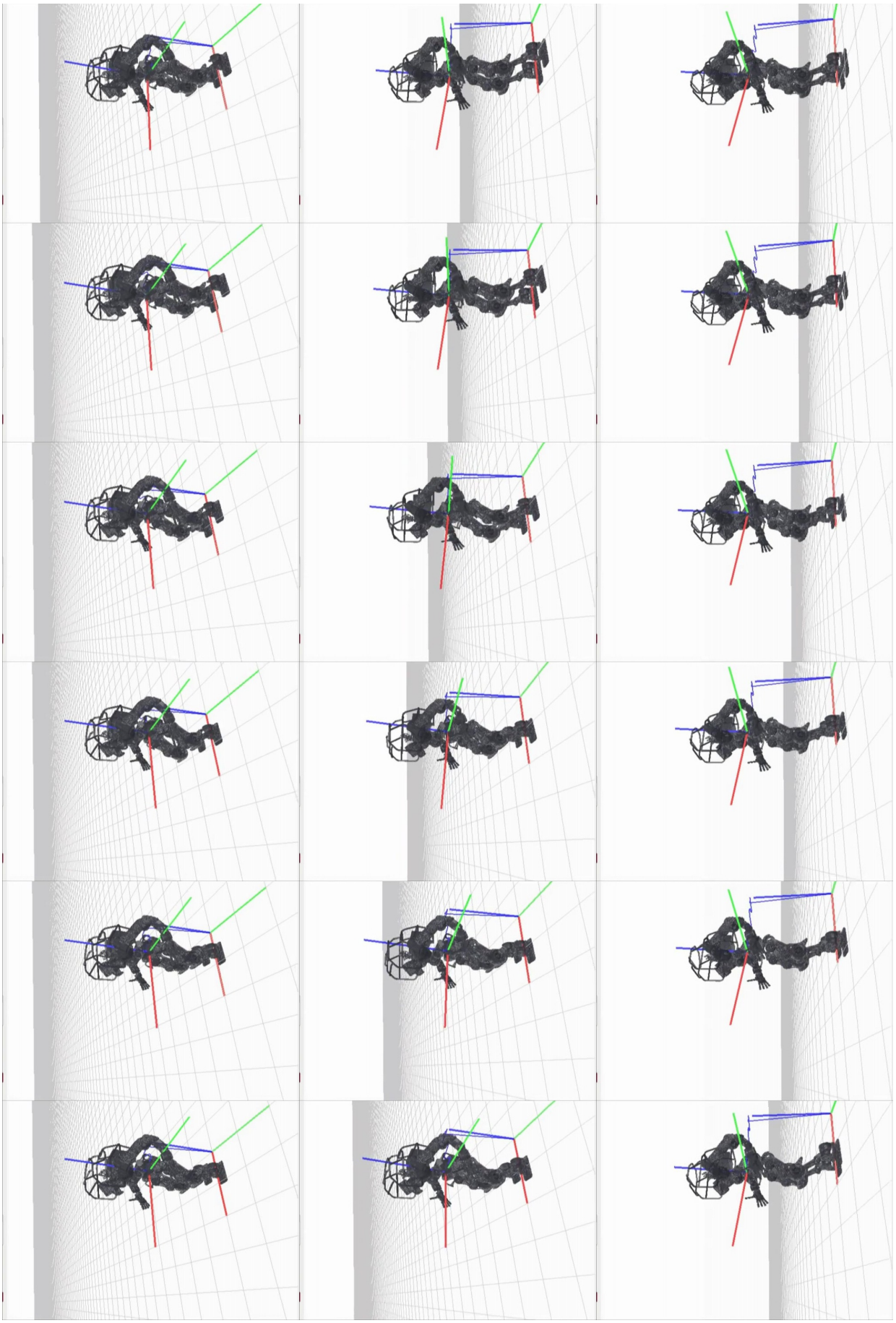


FIGURE 4.6: WALK-MAN simulation experiment: kinematics and F/T sensors.

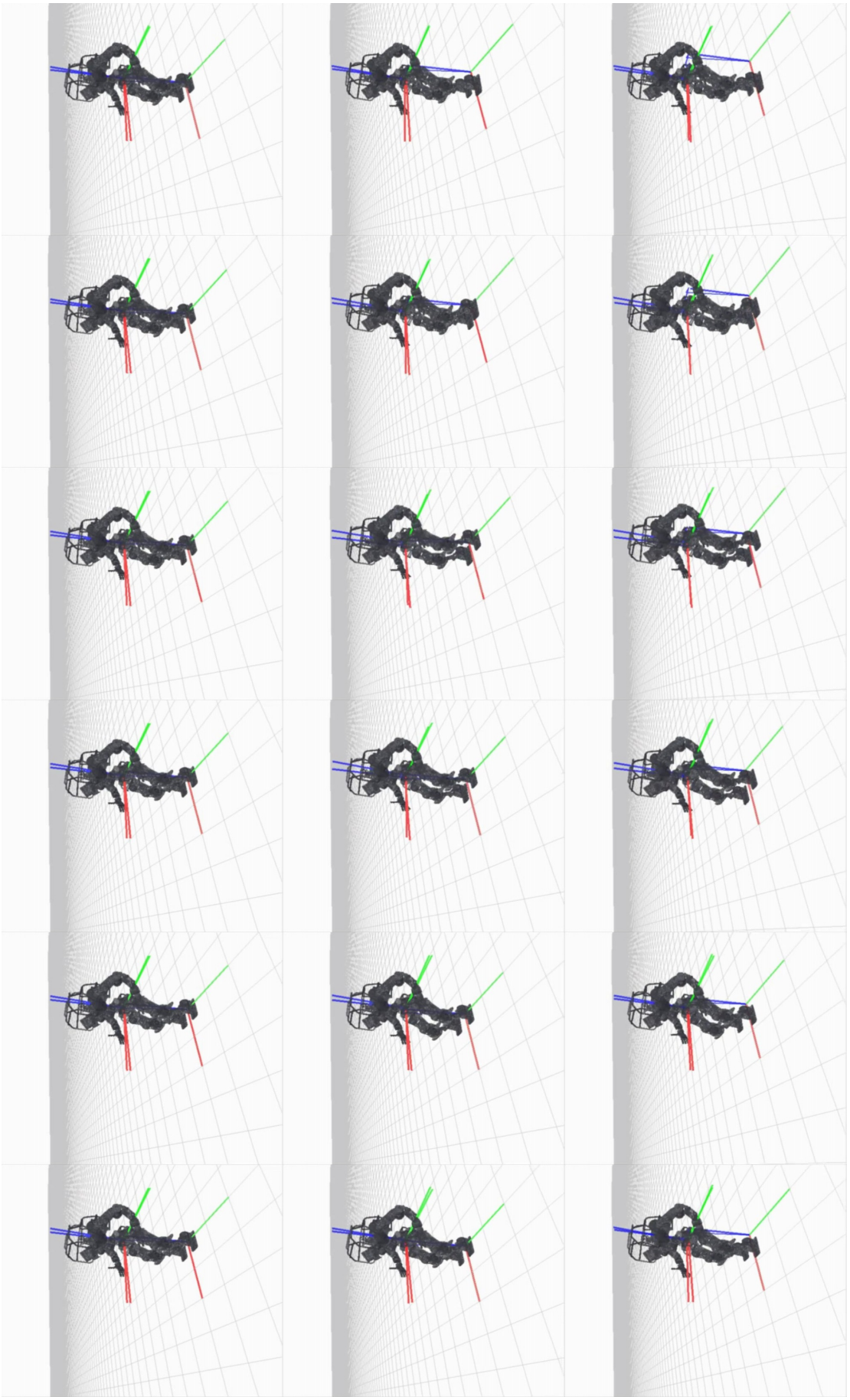


FIGURE 4.7: WALK-MAN simulation experiment: all sensors (1/2).

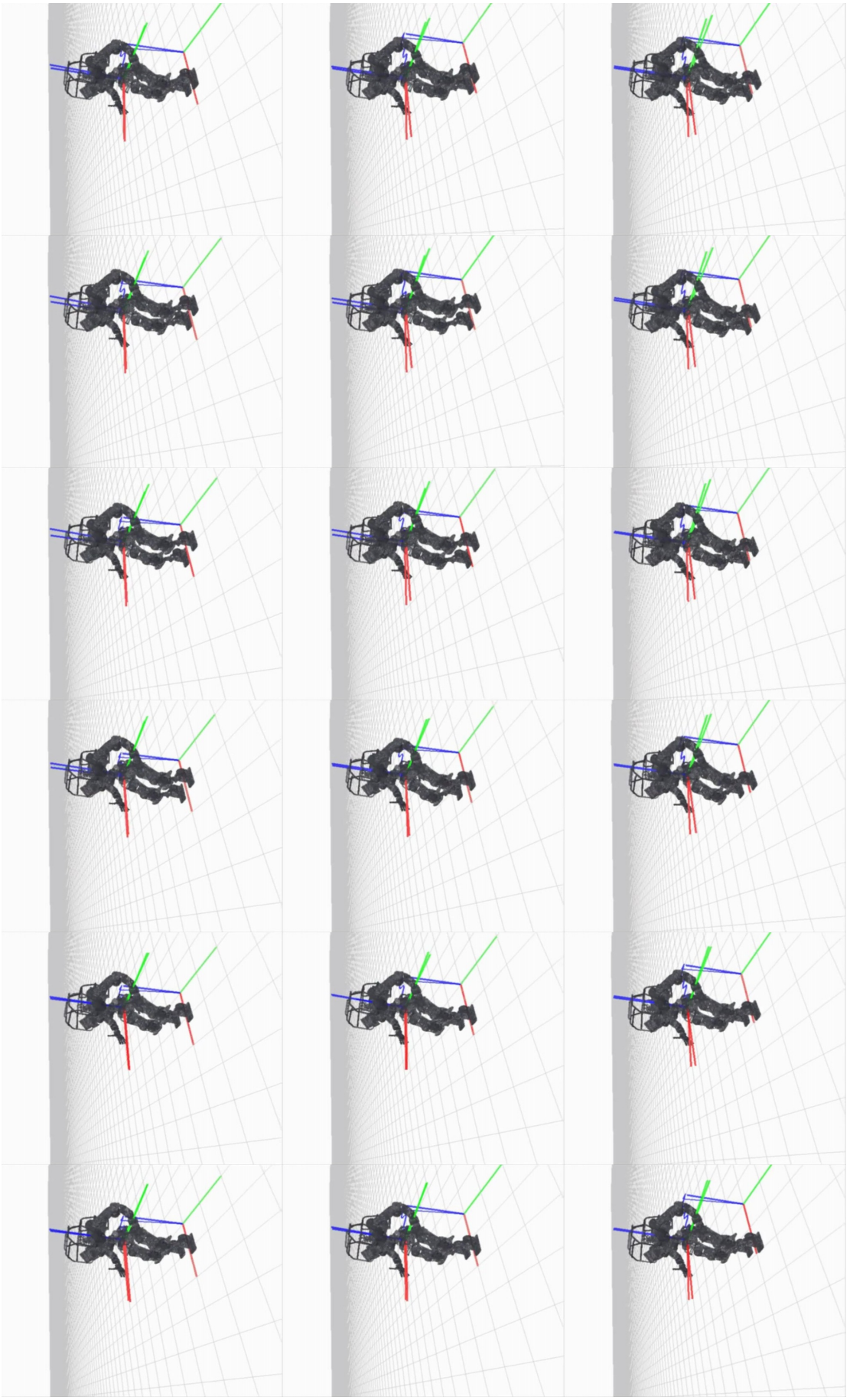


FIGURE 4.8: WALK-MAN simulation experiment: all (2/2).

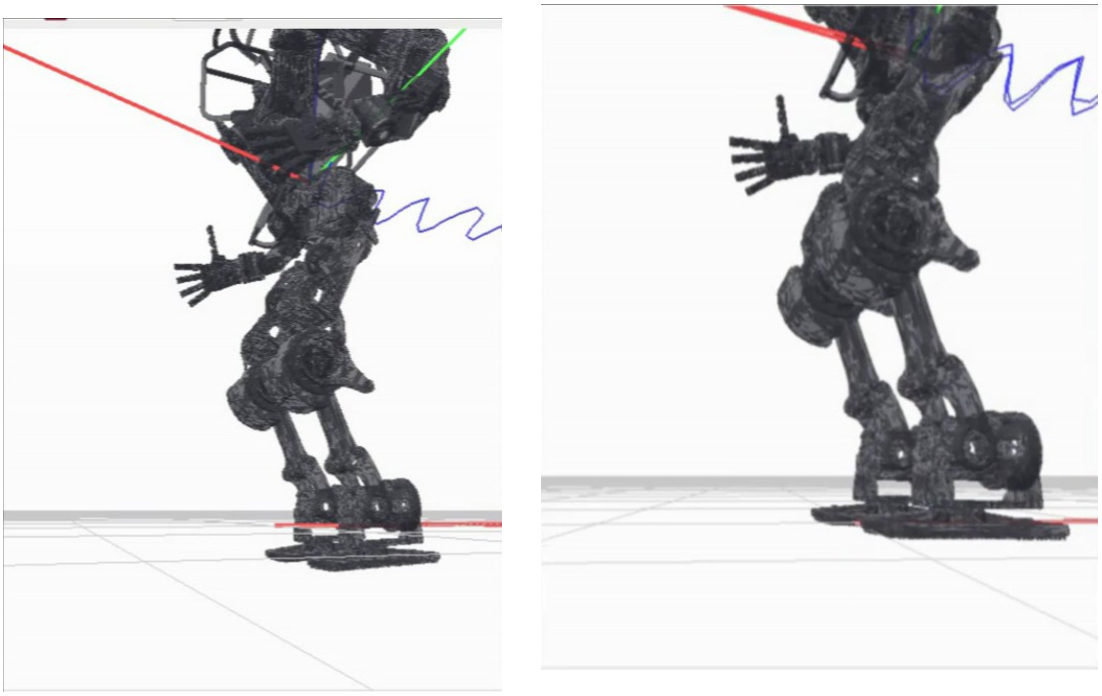


FIGURE 4.9: WALKMAN position estimation without IMU (left) and with IMU (right). As the IMU provides information about floor inclination, without it the robot will either sink (as in this case) or start floating, whereas with it the robot should remain in contact with the ground.

Having confirmed the above, we used data recorded in previous experiments to test the proposed fusion algorithm, comparing it with a kinematics only based algorithm [100]. For the proposed fusion algorithm, three sources of inertial data were used: an LPMS IMU, a VectorNav VN-100 IMU (Fig. 4.10), and an ideal IMU which measured a constant downward gravity vector. The LPMS and VectorNav IMUs were mounted on the robot's base link, whereas the ideal IMU data was artificially created. Results of this experiment can be seen on figures 4.11, 4.12 and 4.13.

4.5.2 Walking Experiments

We performed walking experiments with the robot. We used a motion capture system comprised of five cameras, and placed four markers on WALK-MAN's base link, as can be seen in Figure 4.14, to obtain the ground truth of that link's position and orientation. Results of this experiment can be seen on figures 4.16, 4.17 and 4.18.

TABLE 4.1: VN-100 IMU parameters

Model	VN-100
Sampling rate	(400, 800) [Hz]
Resolution	$< 0.05[deg]$
Accelerometer	3 axis $\pm 160[m/s^2]$
Gyroscope	3 axis $\pm 2000[deg/s]$
Magnetometer	3 axis $\pm 2.5[gauss]$
Pressure sensor	$10to1200[mbar]$

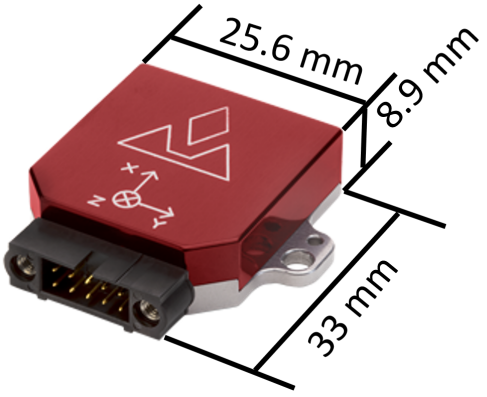


FIGURE 4.10: Used VN-100 IMU from VectorNav ©. (From [101]).

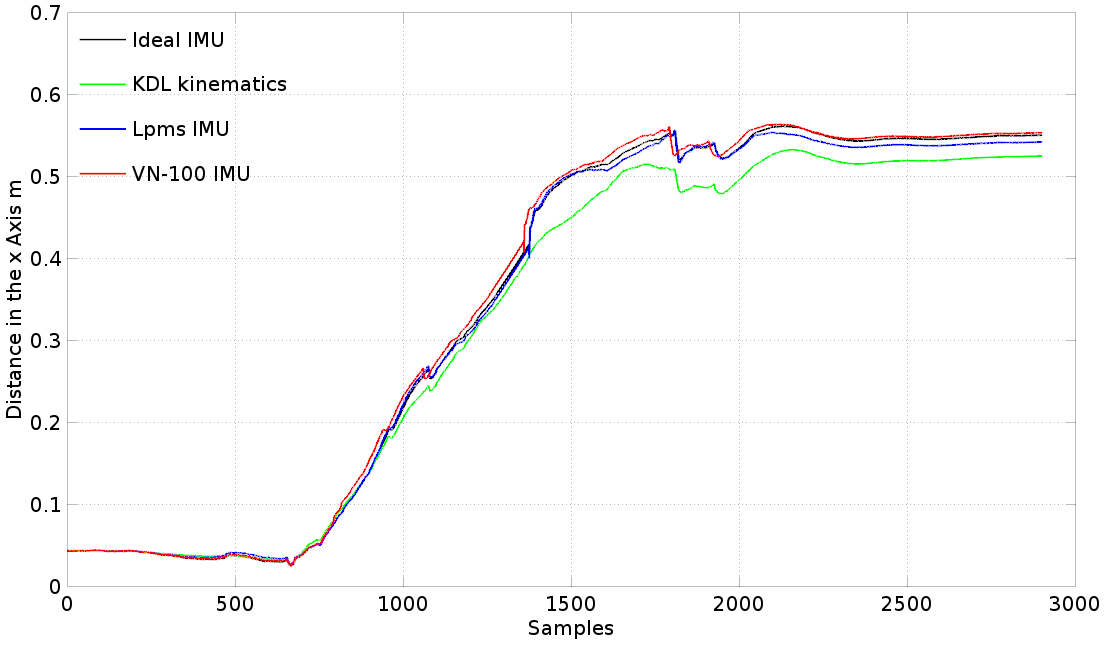


FIGURE 4.11: Comparison on the x axis for the walking experiment simulation with WALK-MAN. Traveled distance in the vertical axis versus samples in the horizontal axis.

4.6 Discussion

We will organize the discussion of this chapter according to the objectives set in the beginning.

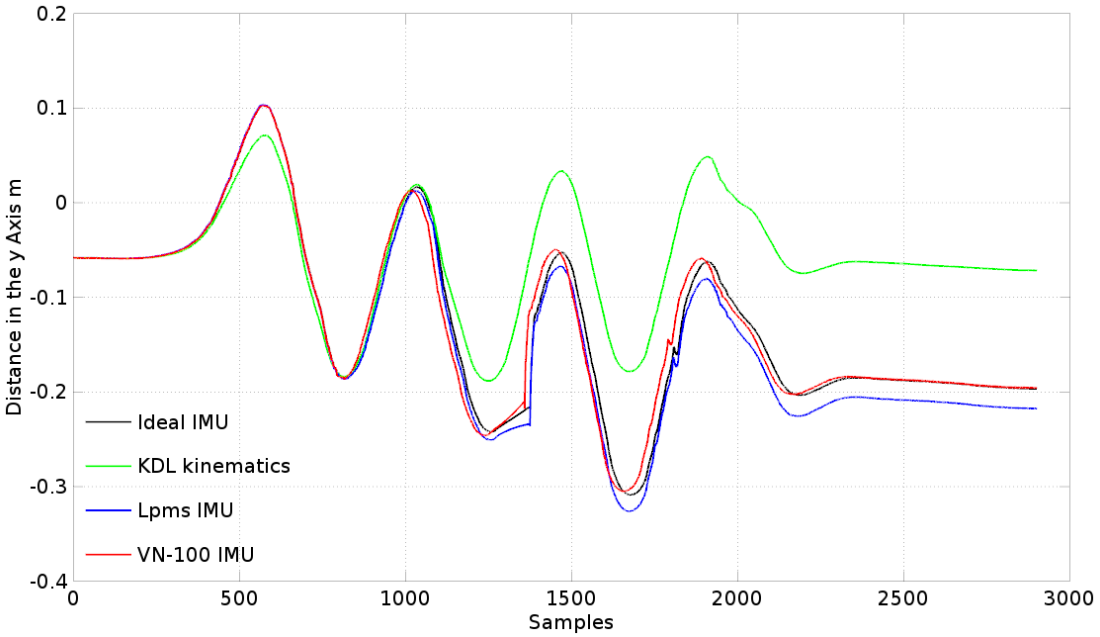


FIGURE 4.12: Comparison on the in y axis for the walking experiment simulation with WALK-MAN. Traveled distance in the vertical axis versus samples in the horizontal axis.

4.6.1 Body Localization System Design

We designed a sensor system with similar inputs as those used for the body position and orientation estimation by humans, that is, motor enconders and 6 axis force/torque sensors as proprioception, and an IMU mounted on the body of the robot as somatic graviception. The IMU was used to know the orientation of the base link of the robot, while the motor encoders were used to obtain the kinematic configuration of the robot. Then the force/torque sensors were used to decide the foot in contact with the ground, and knowing this, forward kinematics were used to estimate the position of the base link of the robot. We chose an EKF based sensor fusion algorithm to combine the information from the forward kinematics with the IMU data, to obtain a better estimate.

As a future work it remains to test other sensors modalities that humans have, such as the skin, which could be a very good source of information, and for which very promising artificial versions for robots are been developed [102, 103].

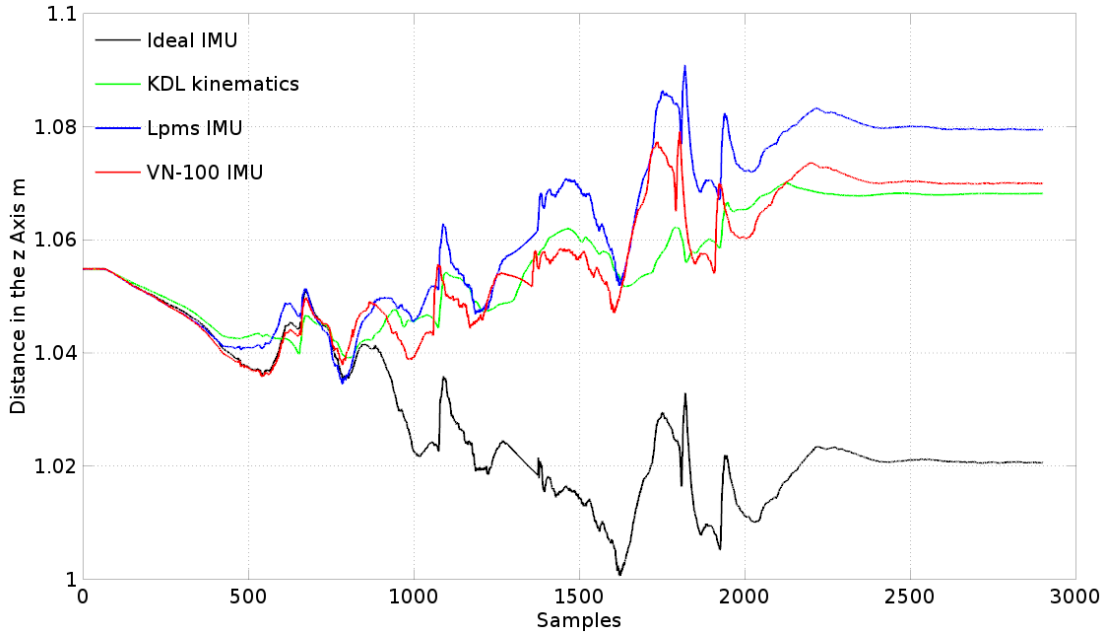


FIGURE 4.13: Comparison on the in z axis for the walking experiment simulation with WALK-MAN. Traveled distance in the vertical axis versus samples in the horizontal axis.

4.6.2 Evaluation of the System in Simulation

We confirmed, through simulation experiments using logged data from actual walking experiments of the robot, that the fusion of data from joint encoders, IMU and force/-torque sensors on the feet effectively improves the position and orientation estimation. For the fusion algorithm, we compared the chosen algorithm with a kinematics only based algorithm [100], and confirmed that the IMU is compensating for errors given by the inclination of the floor, which cannot be completely sensed only with kinematics and force/torque information. Using either the LPMS IMU or the VectorNav VN-100 IMU did not affect significantly the estimation, i.e., both IMUs showed good performance. Care must be taken on the tuning of the parameters of the Schmitt trigger used to decide when and which foot is in contact with the ground, as all the estimations depend on this information.

4.6.3 Walking Experiments

We performed a series of simple walking experiments with the humanoid robot WALK-MAN, using a motion capture system to obtain the ground truth of the walking trajectory of the robot. We achieved a performance comparable to that obtained using the

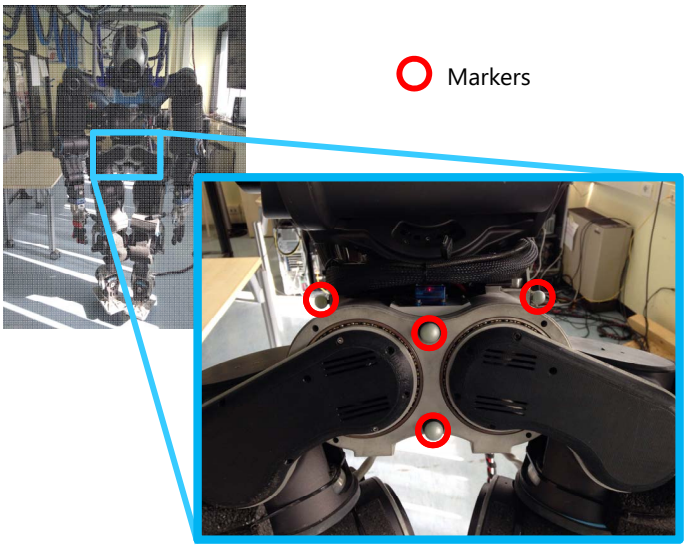


FIGURE 4.14: Marker placement on WALK-MAN for walking experiments with motion capture.

same algorithm on the humanoid robot Atlas, from Boston Dynamics [98]. Once again, care must be taken on the tuning of the parameters of the Schmitt trigger for ground contact classification. Other approaches to classify the ground contacts could also be used, such as using IMU's also on the feet or other links, despite loosing the human-likeness of the system, or using robotic skin sensor as already mentioned. The success of the implementation of the kinematic localization algorithm, originally designed for the humanoid robot Atlas, on the humanoid robot WALKMAN, showed the platform independence of this algorithm, opening the possibility to also implement it on the robot used in our laboratory, WABIAN-2R.

4.7 Summary

In this chapter we proposed a Body Localization System, using information from the joint encoders, an IMU on the base link of the robot, and force/torque sensors on the feet, for the estimation of the position and orientation of the robot's CoM. A loosely coupled EKF based algorithm was chosen to combine the information form the different sensors.

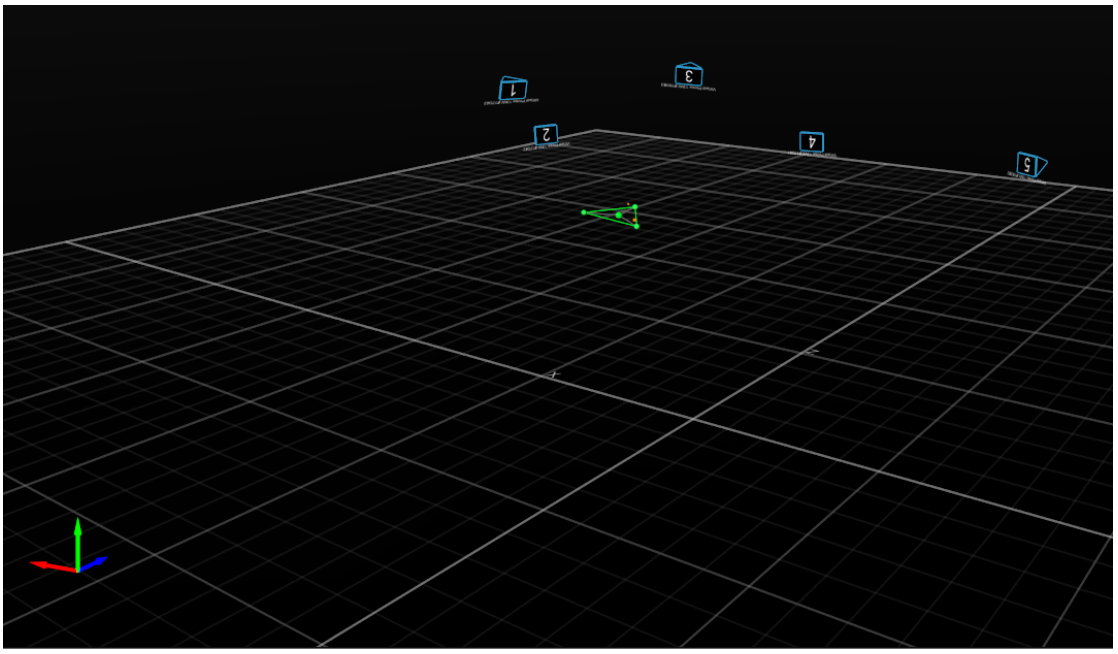


FIGURE 4.15: Motion capture environment for walking experiments with WALK-MAN.

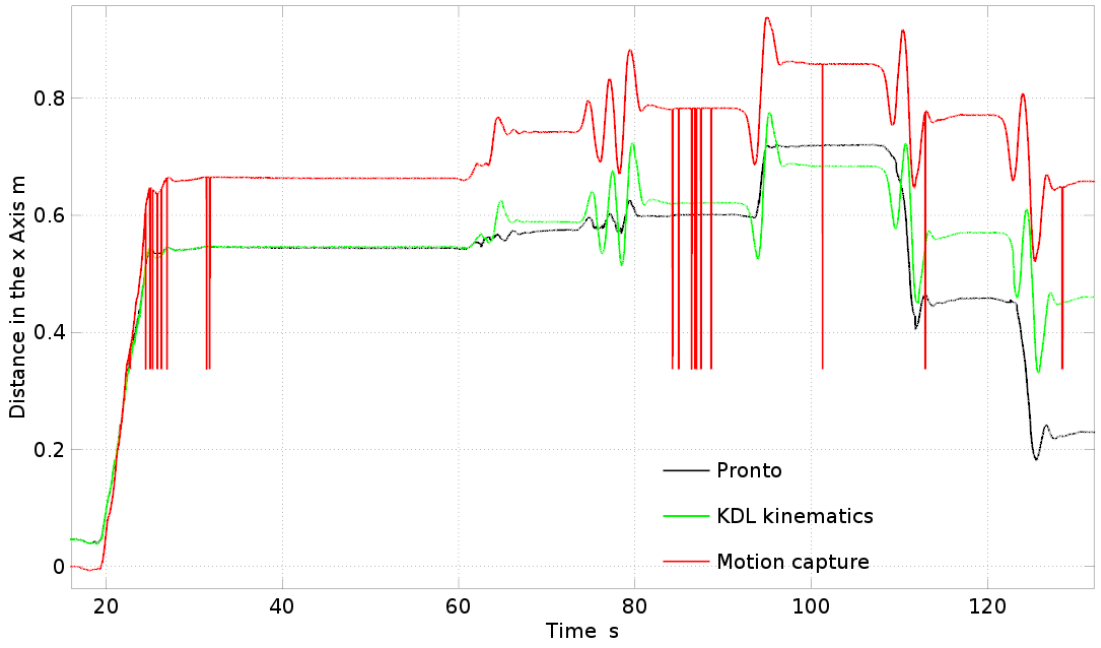


FIGURE 4.16: Comparison on the in x axis for the walking experiment with a motion capture system as ground truth. Traveled distance in the vertical axis versus time in seconds in the horizontal axis.

Tests of the chosen algorithm were performed in a simulation environment, using data from actual experiments of the robot, and finally walking tests were performed using a motion capture system to compare the estimated trajectories with a ground truth.

Having this system , in the next chapter we will explore how the walking motions affect

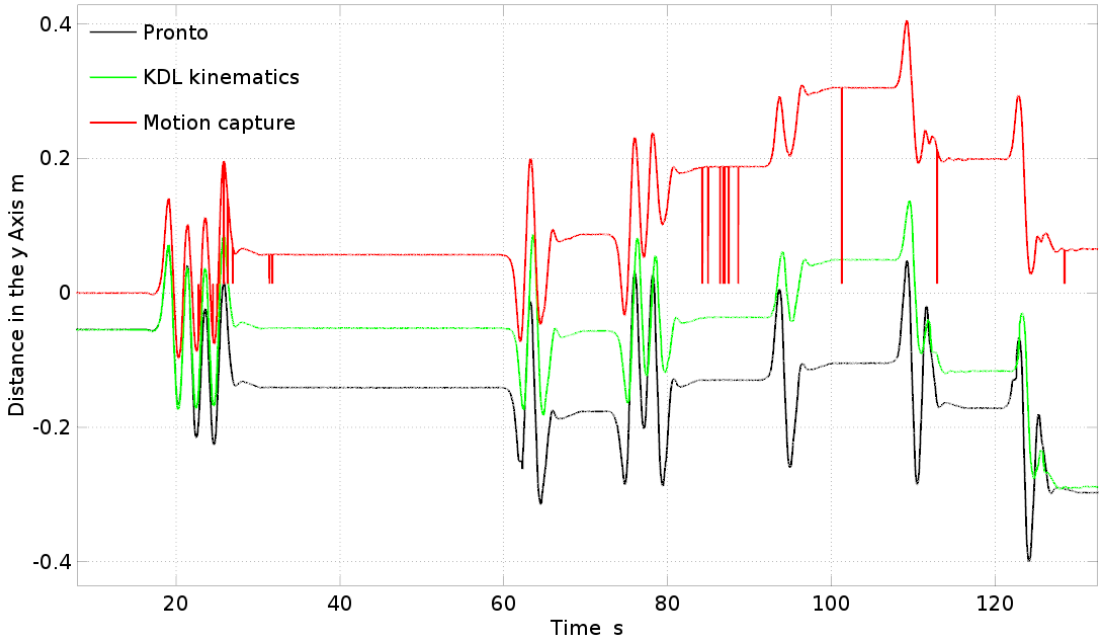


FIGURE 4.17: Comparison on the in y axis for the walking experiment with a motion capture system as ground truth. Traveled distance in the vertical axis versus time in seconds in the horizontal axis.

the performance of the proposed localization algorithm, as we did in Chapter 3 for visual localization algorithms.

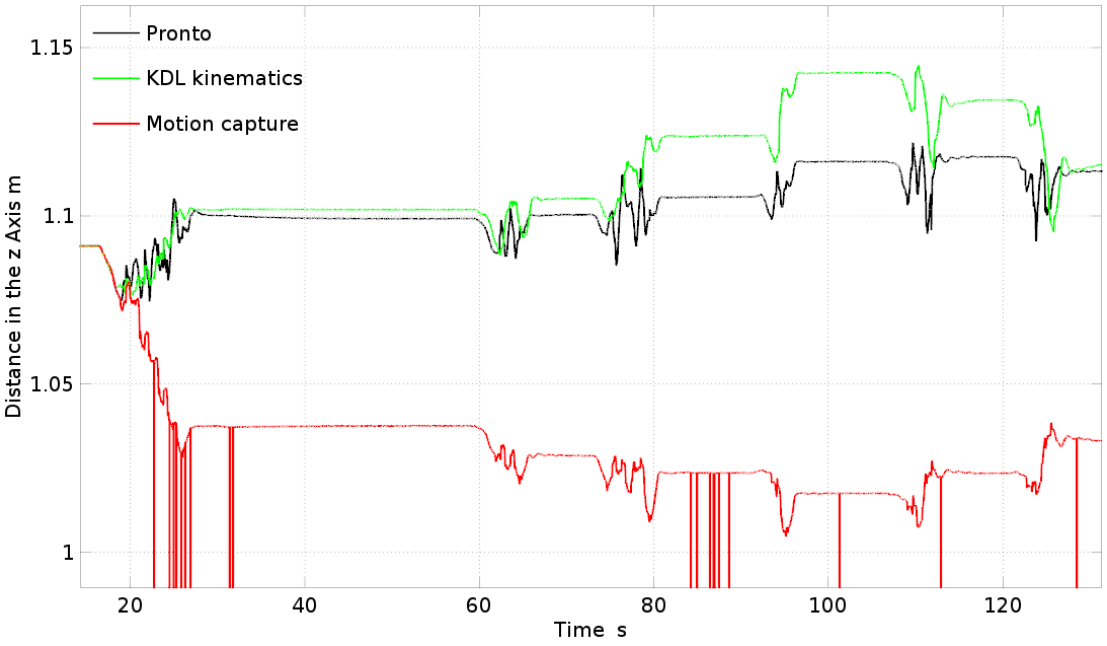


FIGURE 4.18: Comparison on the in z axis for the walking experiment with a motion capture system as ground truth. Traveled distance in the vertical axis versus time in seconds in the horizontal axis.

Chapter 5

Effects of Walking on Body Localization

5.1 Introduction

In Chapter 3 we explored the effects of walking gait parameters as step length or walking symmetry on the localization performance of a humanoid robot using the sensor module developed in Chapter 2, and we found that the localization errors using visual algorithms indeed changed depending on these walking parameters.

In this chapter, we assess the effects of walking parameters on the localization of the robot but this time using kinematic odometry.

We already pointed out the importance of having the ability to self-localize in the environment, specifically for humanoid robots to become more useful in our daily lives. Also, we mentioned that a common ways for robots to self-localize is through odometry algorithms, and in Chapter 3 we focused on VO algorithms. In this chapter, we will focus on the effects of walking parameters on the estimation of the robot's change in position through the use of motions sensors and its kinematics.

We mentioned active localization as a way to improve the localization performance of a robot, and we explored the effects of walking parameters on the localization performance of VO based algorithms.

Moreover, locomotion parameters have effects on self-localization in humans not only when estimating traveled distances through the use of optic flow [60], but also they are known to regulate walking speed in order to improve localization with our eyes closed [87].

With the above in mind, the contribution of this chapter is to answer the following questions regarding kinematic localization systems for biped humanoid robots:

- Does performance of such systems depend consistently and non-trivially with humanoid gait?
- What effects do different walking styles have on the performance of such systems?

The approach in this chapter is data-driven just as in Chapter 3, i.e., we do not try to predict localization performance from simplified mechanical, control, sensor, or environment models. We measure the localization performance of the whole system directly, using a motion capture system to obtain the ground-truth data. We performed many experiments varying the robot's walking gait parameters, and for this chapter We focus on the following parameters:

a) Step Length

b) Walking Style and Symmetry

In the following sections related work and the overview of the used system along with the method to evaluate it and analyze the results will be presented. After that, the performed experiments will be described, and the obtained results presented. Finally I will discuss these results, from the point of view of the related work and human studies.

5.2 Related Work

We presented related works regarding self-localization for humanoid robots using visual odometry algorithms in Chapter 3. However, we can also estimate the ego-motion of humanoid robots through the kinematics and dynamics of the robot.

A Kalman filter based CoM state estimator using linear inverted pendulum dynamics as the process dynamics is proposed in [104]. Various process and output models are constructed and compared using a force controlled humanoid robot. They use three different sensing models, position, center of pressure and total CoM force, and build four different estimators, one naive, one for CoM offset, one for external forces and a dual estimator.

In [105], a CoM and disturbances estimator using only an IMU and forward kinematics is proposed. A Kalman filter based approach is used, and the disturbance is modeled based on the previous CoM and ZMP states, using a quadratic programming method. Simulations are carried out using the model of the humanoid robot SURALP (Sabanci University Research Laboratory Platform) [106].

Another CoM motion estimator based on a complementary filter that combines kinematic and acceleration information is proposed in [107]. They assume that the contact point moves with respect to the ground at the instantaneous minimum velocity point, which they estimate by optimization and use to improve the accuracy of the kinematics. Then in [108] they propose a Kalman based approach fusing kinematic information, the double integral of CoM acceleration, and the relationship between CloM and ZMP. Both works are tested in the simulation platform OpenHRP3 [109].

Many of these estimators are based on Bayesian approaches. For instance, Xinjilefu et al. [110] propose a decoupled estimation, first using a joint dynamics estimator, and then a base link position estimator, instead of including all that information in a single filter, in order to reduce the computational cost sacrificing some accuracy. Then, in [111], a bipedal robot state estimator is proposed, based on another originally designed for a quadruped robot [112]. These estimators make the filter update based on feet measurements.

In [113] an approach to estimate the dynamic pose, as well as the internal and external wrenches acting on on the individual feet of a bipedal robot is presented. They achieve this by fusing haptic, inertial and force/torque measurements obtained from a compliant tactile sensor array on the feet soles, IMU's and F/T sensors on the robot's ankles, respectively, and using an EKF based approach with state augmentation.

Using the biped robot LOLA [114], a state estimation using multibody dynamics is proposed in [115]. This estimator is also based on a Kalman filter, that uses the dynamical model of a linear inverted pendulum, and fuses sensor information from motor encoders, IMUs and F/T sensors. It also uses a feed-forward computed disturbance input.

From the biological point of view, humans mainly use visual, gravitational/inertial and proprioceptive cues for ego-motion estimation. Moreover, humans change the weight they put on each sensory input depending on the situation [17], [83]. Also, we know humans plan their walking gait ahead in many situations, as to keep stability in difficult situations like slippery terrains [84], but we also change our gait parameters when there are problems with the sensory inputs, as decreasing walking speed or having a more backward leaning trunk posture when visual disturbances arise [85]. Moreover, different sensory modalities perform better depending on our motions. The visual system performs better at lower frequencies than the vestibular system, but both are integrated in an optimal manner [86]. There is also evidence pointing out that modifying the walking speed has effects on our path integration abilities, making us overestimate distances when walking at slower speeds [87], as well as walking cadence affecting the performance of path integration, achieving the best performance at about 2 Hz [88].

5.3 System Overview

For the experiments in this chapter we used the biped humanoid robot WABIAN-2R [53] (Fig. 5.1), a 33 DoF bipedal humanoid robot (details in Section 3.3.1), with an IMU mounted on the lower back

For the ground truth measurements, the motion capture system OptiTrack V120:Trio at 120 fps was used, placing the photo-reflective markers on the camera to obtain the actual trajectory.

The different reference frames and transformations used for the experiments can be seen on Fig. 5.3. We use three main reference frames, the World frame, C_t, the frame of the camera system at time t, and F_t, the frame of the contact foot at time t. Also, following the notation used in [89], we define ${}^{(est)}T_{A_{t_i} \to B_{t_j}}$ as the transformation of frame B at time t_j relative to frame A at time t_i, calculated with the estimator est. The motion capture system tracks the head motion in the world frame, ${}^{(gt)}T_{W \to C_t}$, whereas the kinematic

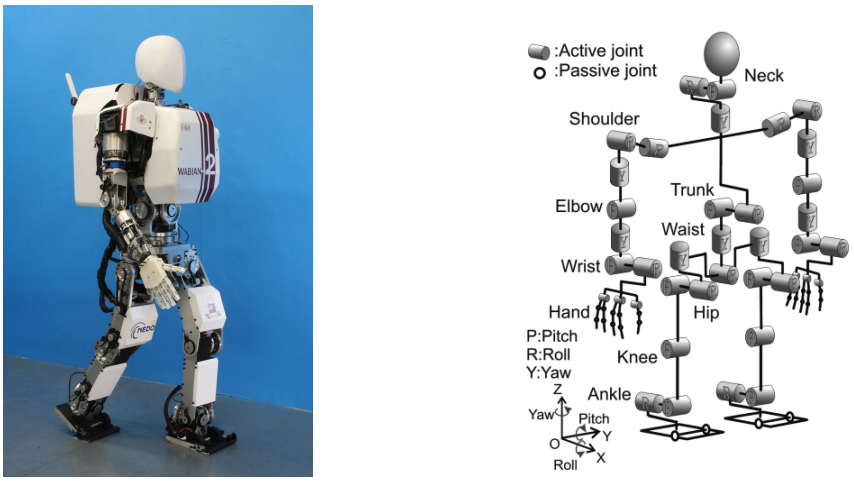


FIGURE 5.1: Robotic platform WABIAN-2R (left) and DoF configuration (right).

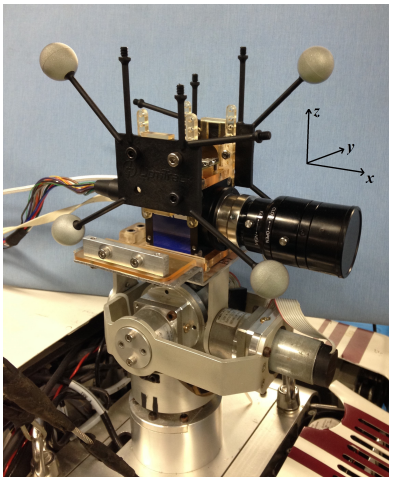


FIGURE 5.2: Close-up of the head system used for localization and ground-truth (head, camera, reflective markers)

odometry estimates the motion of the robot's head relative to the contact foot frame at each time stamp, ${}^{(kin)}T_{F_t \to C_t}$.

For the kinematic localization, we used pronto state estimator, which we presented in Section 4.3 in Chapter 4.

We also logged acceleration and angular velocity data at 200 Hz from one IMU mounted on the camera itself, as well as force and torque data from sensors placed on both feet, also at 200 Hz. This data was processed and analyzed to look for possible differences between different walking speeds (Figs. 5.4, 5.5).

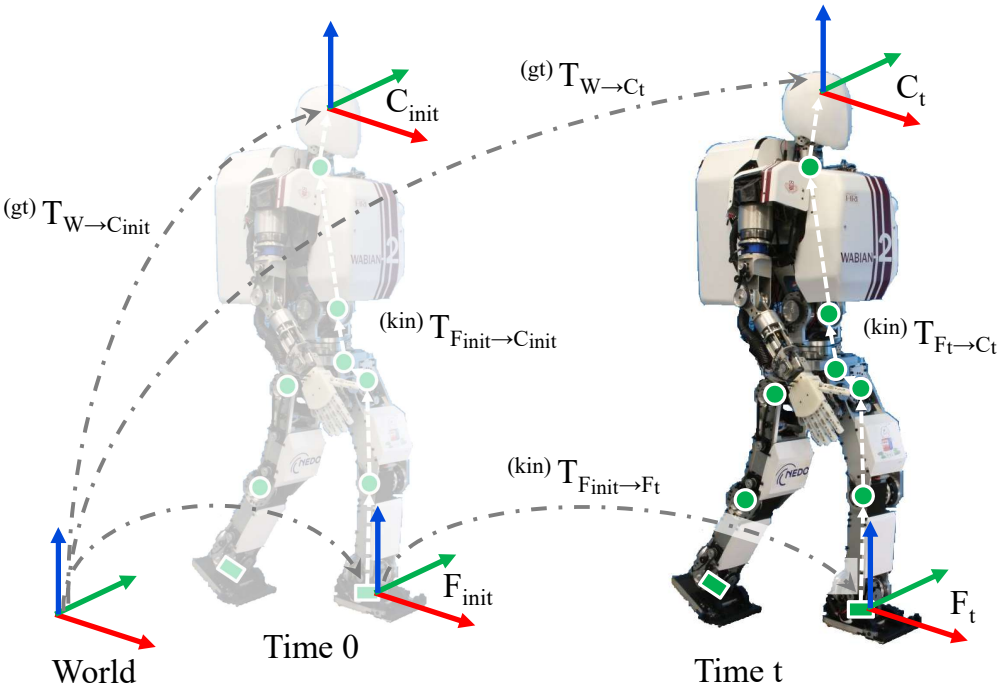


FIGURE 5.3: Used coordinate frames and transformations. C_t is the frame of the camera system at time t. F_t is the frame of the contact foot at time t. The motion capture system tracks the head motion in the world frame, ${}^{(gt)}T_{W\to C_t}$, whereas the kinematic odometry estimates the motion of the robot's head relative to the contact foot frame at each time stamp, ${}^{(kin)}T_{F_t\to C_t}$.

5.4 Data Analysis

We used the same metrics as those presented in section 3.4, namely the absolute trajectory error for the estimated trajectory's overall error, and the relative pose error to assess the estimated trajectory's drift w.r.t. the ground-truth [93]. The equations for these metrics are presented below.

$$ATE_t = {}^{(gt)}\mathrm{T}^{-1}_{W\to C_t}\,{}^{(vo)}\mathrm{T}_{W\to C_t} \tag{5.1}$$

$$RMSE(ATE_t) = \left(\frac{1}{n}\sum_{i=1}^{n}\|ATE_i\|^2\right)^{\frac{1}{2}} \tag{5.2}$$

$$RPE_t = {}^{(gt)}\mathrm{T}^{-1}_{C_t\to C_{t+\Delta}}\,{}^{(vo)}\mathrm{T}_{C_t\to C_{t+\Delta}} \tag{5.3}$$

$$RMSE(RPE_t) = \left(\frac{1}{m}\sum_{i=1}^{m}\|RPE_i\|^2\right)^{\frac{1}{2}} \tag{5.4}$$

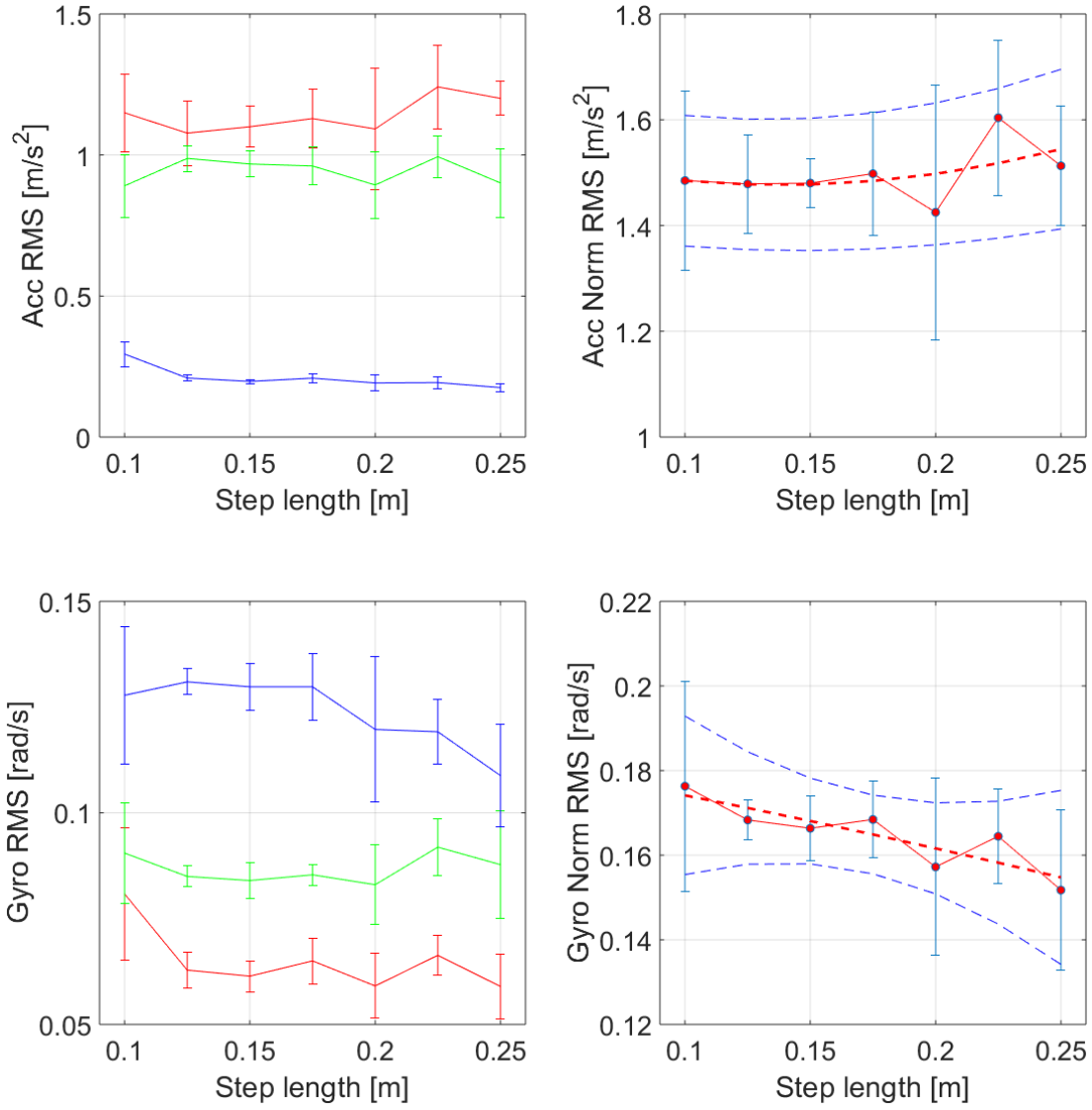


FIGURE 5.4: RMS of the data from the accelerometer and gyroscope of the IMU mounted on the camera for x (red), y (green) and z (blue) (right), and their norm, with a fitted quadratic curve (left). The used step length is on the horizontal axis.

5.5 Walking Speed Experiments

As explained in Section 5.1, in this chapter we focus on the effects of step length, and hence walking speed, on localization performance. We generated one walking pattern for each step length, all for which the total walking distance was fixed to 1.5 m on a straight line. The used step lengths were 0.1, 0.125, 0.15, 0.175, 0.2, 0.225 and 0.25 m. Step lengths shorter than 0.1 m resulted in unstable gaits, while step lengths longer than 0.25 m were not tested since they were close to the mechanical limits of the robot. The step width was maintained constant at 0.08 m. Five runs were performed for each step length with the robot having the knees bent, i.e., maintaining a fixed height for the

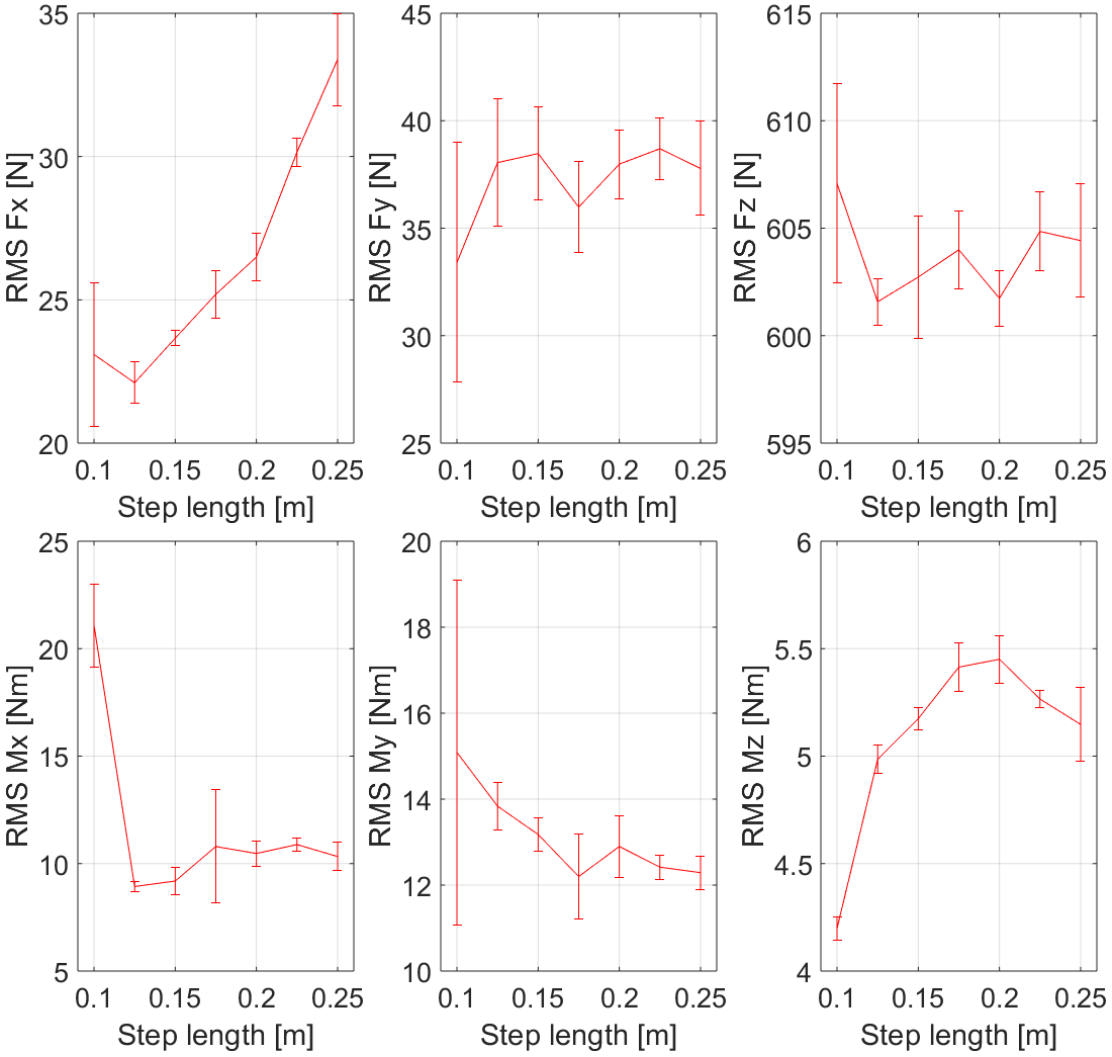


FIGURE 5.5: RMS of the data from the F/T sensors on the robot's feet. The used step length is on the horizontal axis.

center of mass (CoM), and therefore for the camera. The reference walking cadence was fixed to 0.96 s/step, 0.06 seconds for double support phase and 0.9 seconds for single support phase. All patterns were executed on the robot by joint position control without any state estimation (i.e. assuming the reference trajectory of the base was executed perfectly). The motion capture and robot's joints, force, IMU and image data were stored and later analyzed.

Results

An interesting relationship between the kinematic localization accuracy and the robot's step lengths used to cover the 1.5 m trajectory emerged from the results from the walking

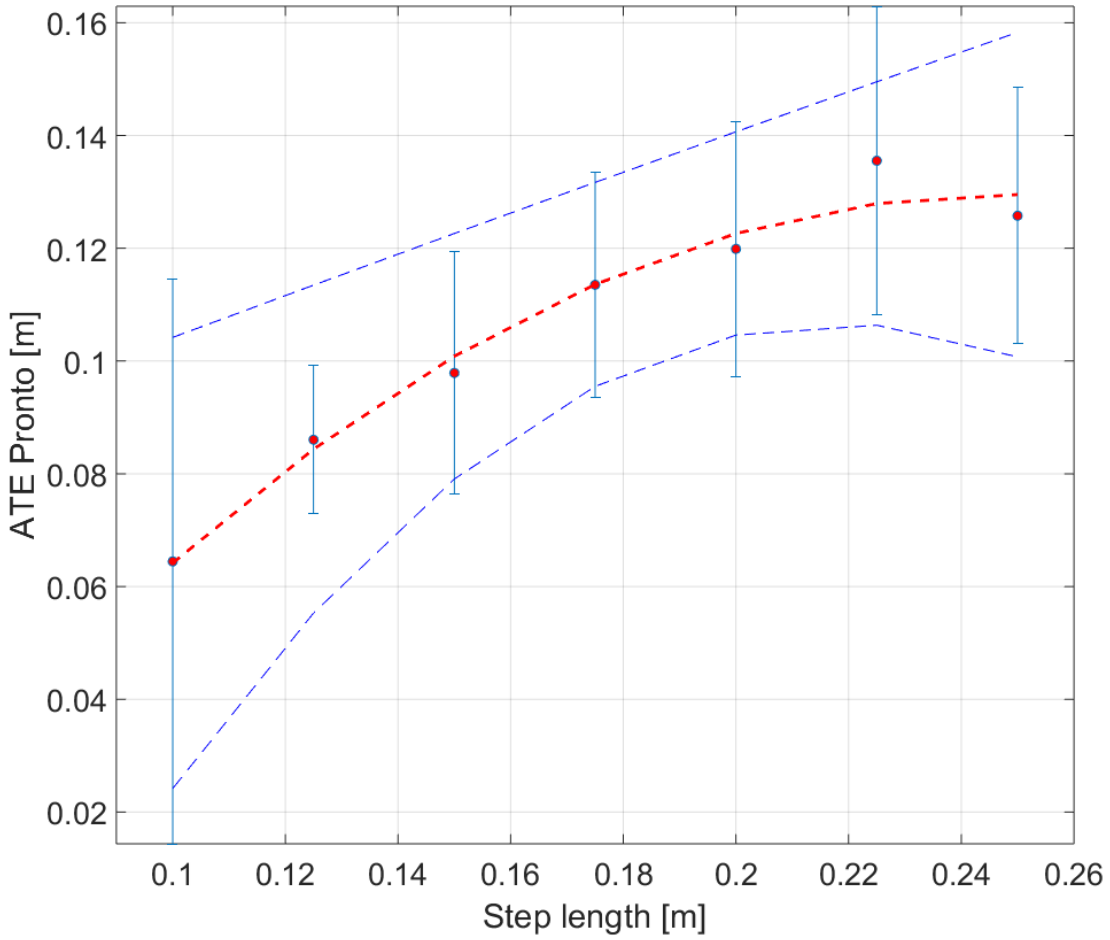


FIGURE 5.6: ATE versus step length for the kinematic odometry. Red dots with blue vertical error bars denote the average and standard deviations for each step length, while the red dashed lines are the fitted quadratic curves for the averages. Fitted quadratic curves were calculated using the *polyfit* function of MATLAB ®.

experiments. The estimation from the kinematic localization algorithm was affected by accelerations, possibly vibrations caused by the walking motions. As the step length increases, both the acceleration and the ATE increase (Fig. 5.4, 5.6).

To assess if the differences were significant, we also conducted a one-way ANOVA to assess the effect of step length on localization performance for the kinematic odometry. There was significant effect at the $p < 0.05$ level $[F(6, 28) = 16.249, p < 0.001]$. Post hoc comparison using the Tukey HSD test indicated significant differences between the performances with step lengths of 10 and 15 cm ($M = -0.0334, SD = 0.0087$), 10 and 17.5 cm ($M = -0.0491, SD = 0.0087$), 10 and 20 cm ($M = -0.0554, SD = 0.0087$), 10 and 22.5 cm ($M = -0.0711 SD = 0.0087$), 10 and 25 cm ($M = -0.0613, SD = 0.0087$), 12.5 and 20 cm ($M = -0.0339, SD = 0.0087$), 12.5 and 22.5 cm ($M =$

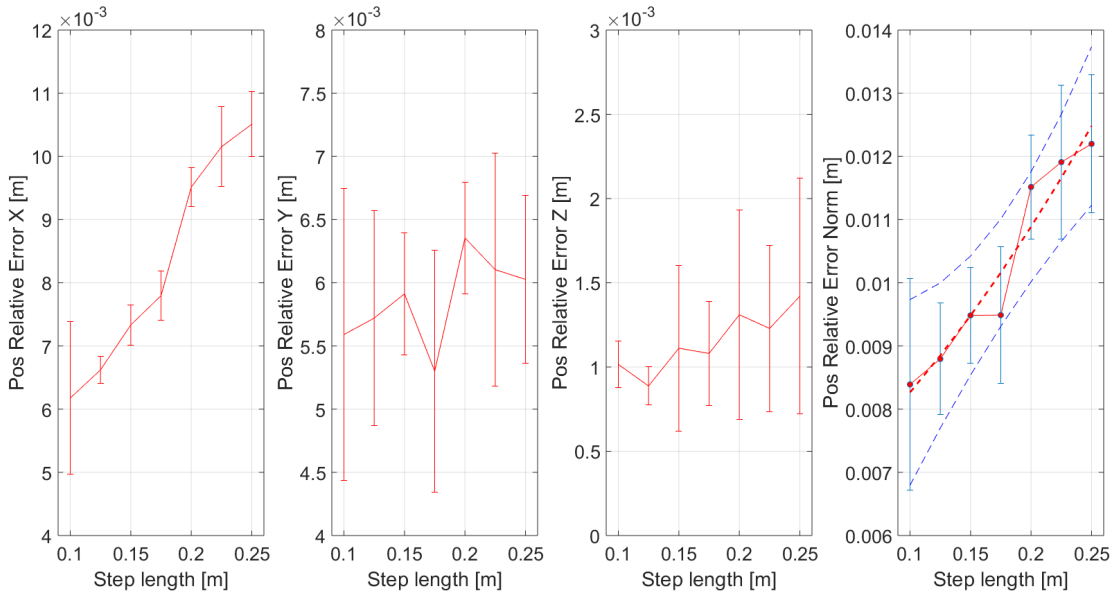


FIGURE 5.7: Relative error of the kinematic odometry algorithm w.r.t. the ground truth. The used step length is on the horizontal axis. The used step length is on the horizontal axis.

$-0.0495, SD = 0.0087$), 12.5 and 25 cm ($M = -0.0394, SD = 0.0087$), 15 and 22.5 cm ($M = -0.0376, SD = 0.0087$) and 15 and 25 cm ($M = -0.0279, SD = 0.0087$).

5.6 Walking Style and Symmetry Experiments

In this section we focus on the effects of walking style and walking symmetry on localization performance. We generated three different walking patterns, one normal walking pattern, one pattern we will call "gallop", and one we will call "slow", which are described in the Section 3.7. For all the patterns, the total walking distance was fixed to 1.5 m on a straight line, and the time to traverse that distance was kept inside the interval between 13.5 and 14.5 seconds. The step width was maintained constant at 0.08 m. Five runs were performed for each pattern. All patterns were executed on the robot by joint position control without any state estimation (i.e. assuming the reference trajectory of the base was executed perfectly). The motion capture and robot's joints, force, IMU and image data were stored and later analyzed.

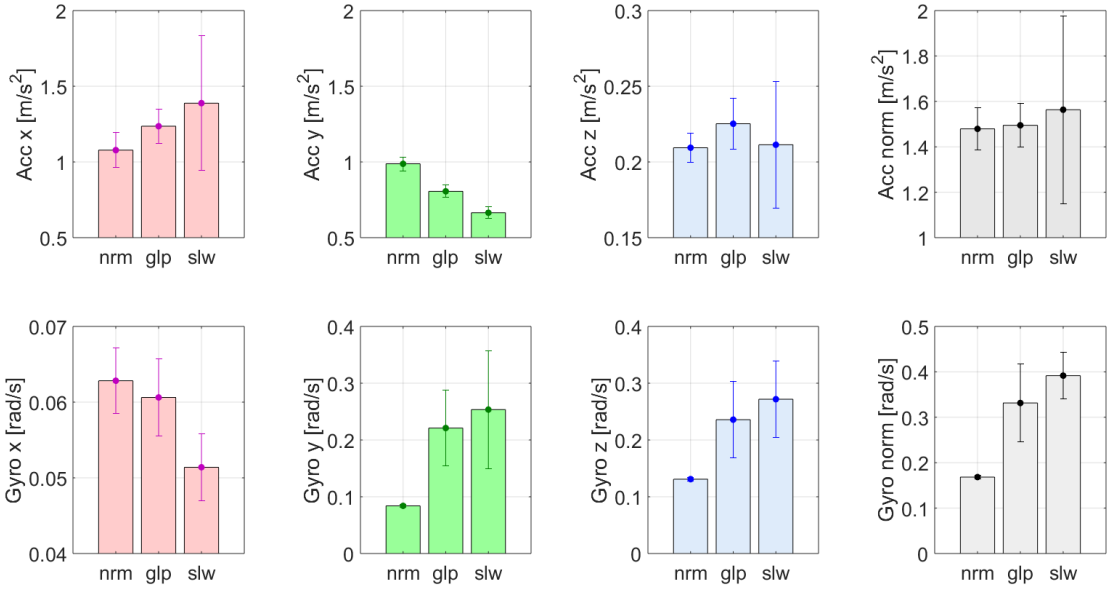


FIGURE 5.8: RMS of the data from the accelerometer and gyroscope of the IMU mounted on the camera for normal (left), gallop (center) and slow (right) in the horizontal axis. Markers with vertical error bars denote the average and standard deviations.

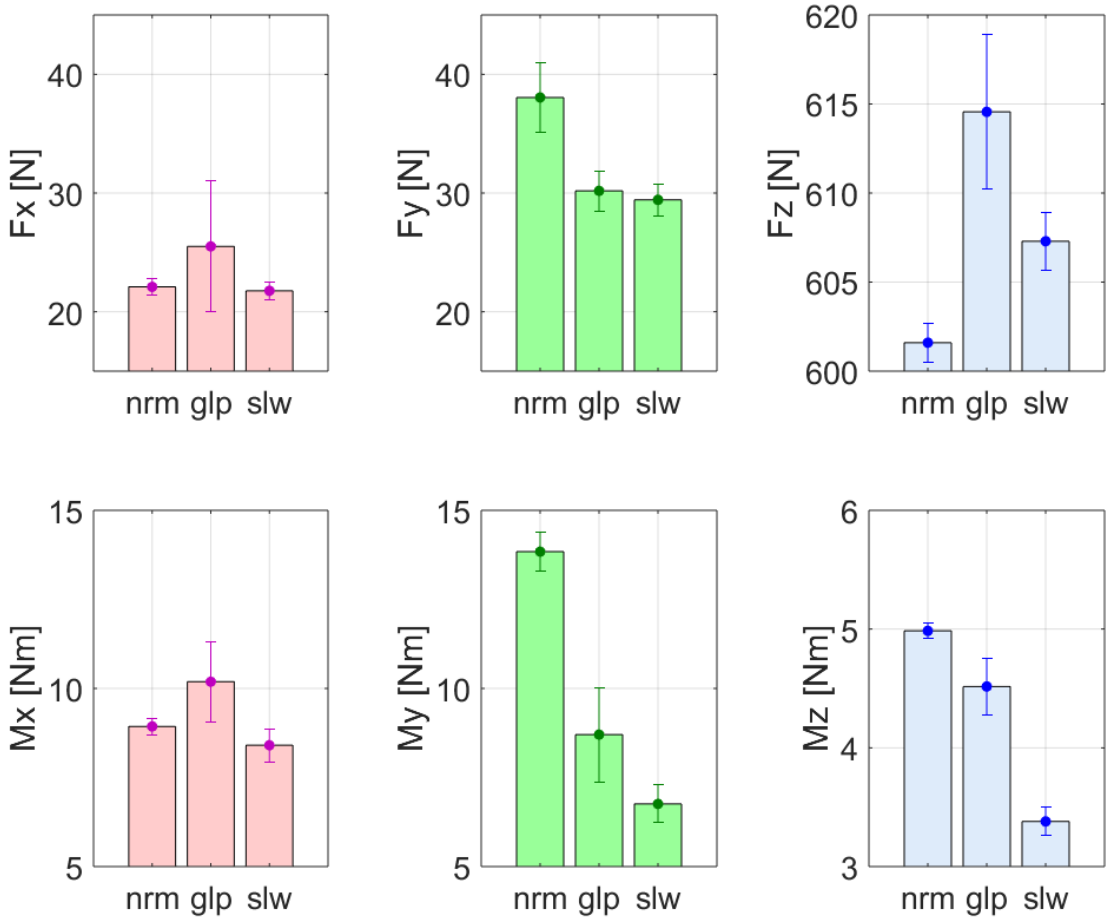


FIGURE 5.9: RMS of the data from the F/T sensors on the robot's feet for normal (left), gallop (center) and slow (right) in the horizontal axis. Markers with vertical error bars denote the average and standard deviations.

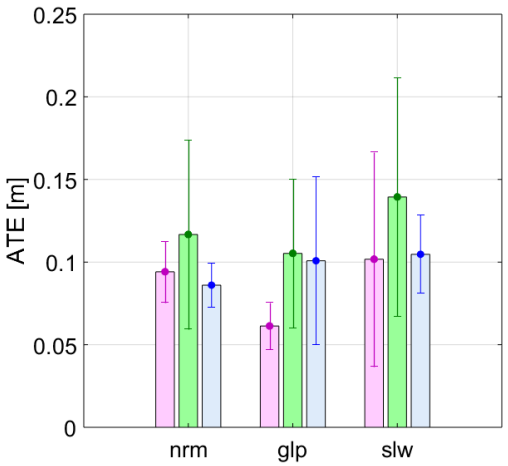


FIGURE 5.10: ATE versus walking styles for SVO 2.0 (left, magenta), ORB-SLAM2 (center, green) and kinematic odometry (right, blue). Walking styles in the horizontal axis are normal (left collection), gallop (middle collection) and slow (right collection). Markers with vertical error bars denote the average and standard deviations.

Results

For the Body Localization System we found that changing both the style and symmetry increased the localization error slightly, as can be seen on Fig. 5.10, which could be explained by the fact that both "gallop" and "slow" gait patterns suffered more reaction forces on the vertical axis than normal walking, as can be seen in Fig. 5.9. Perhaps the reason behind this are bigger impacts while walking with these gaits, which affect both the readings from the joint encoders as well as the monitoring of the contact foot switching from the F/T sensors on the feet, which is crucial for the localization using kinematic odometry. It is also worth mentioning that the kinematic algorithm was the one with the least drift, as can be seen on Fig. 5.11, where the RPE is almost negligible compared to those of the visual algorithms.

A correlation was observed between walking style and localization performance. Therefore the plan is to include these localization performance curves as cost functions within footstep planners [96] such as to minimize localization error. We are also planning to explore other kinds of asymmetries, such as posture asymmetry, for future works.

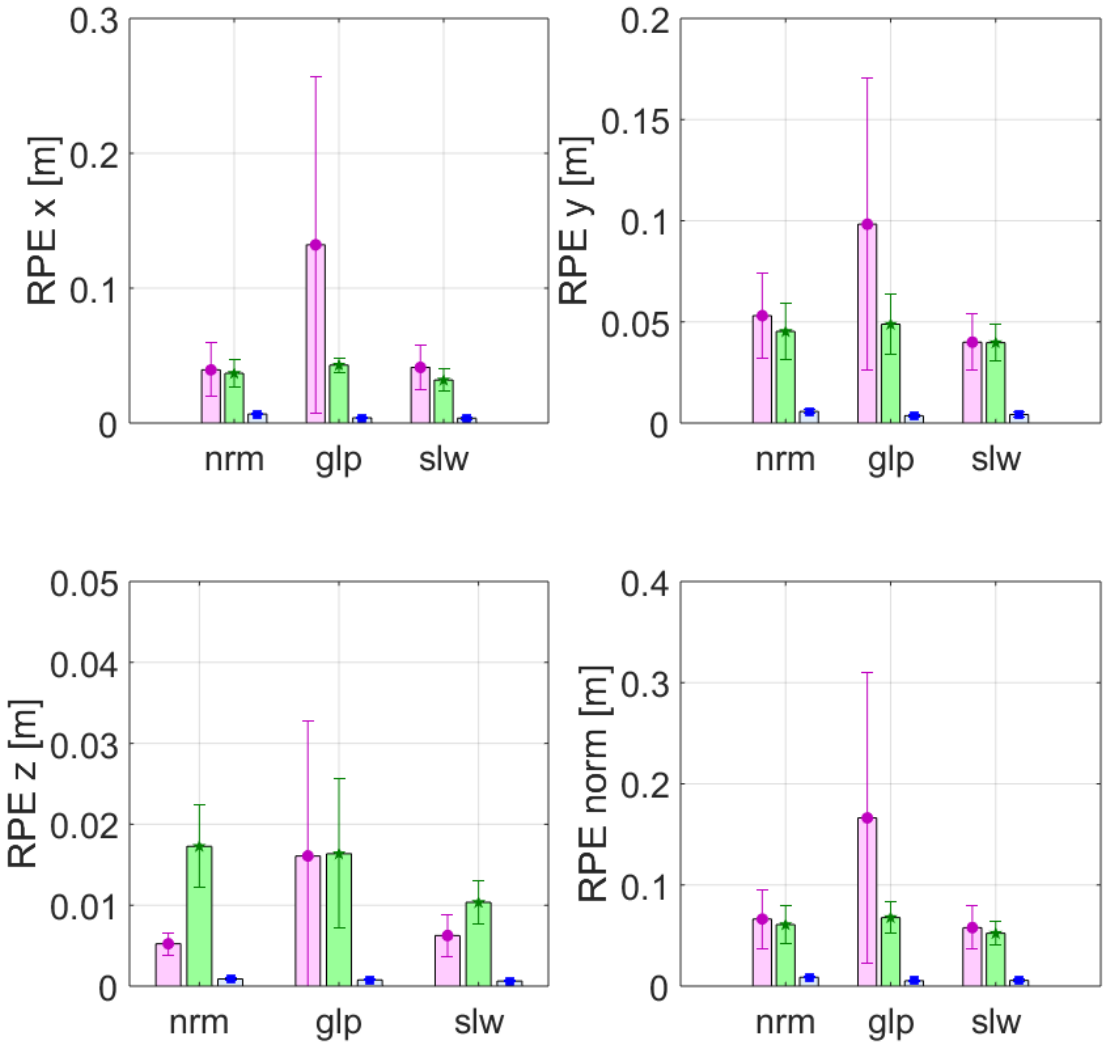


FIGURE 5.11: RPE versus walking styles for SVO 2.0 (left, magenta), ORB-SLAM2 (center, green) and kinematic odometry (right, blue). Walking styles in the horizontal axis are normal (left collection), gallop (middle collection) and slow (right collection). Markers with vertical error bars denote the average and standard deviations.

5.7 Discussion

We will organize the discussion of this chapter according to the objectives set in the beginning.

5.7.1 Walking Speed

As already presented in Section 5.5, we found that kinematic localization performance got worse the faster the robot walked (Fig. 5.6). Our hypothesis is that increased accelerations are affecting this algorithm, as we can see that accelerations also increase

with step length (Fig. 5.4, top). These accelerations could be caused by the increase of vibrations due to bigger impacts of the feet with the ground as the step length increases. Then, these vibrations could be introducing small errors to the readings of the motor encoders, which would accumulate and in the end affect the estimations from the kinematic localization algorithm. Ground reaction forces on the feet in the x axis also increased with step length (Fig. 5.5). These ground reaction forces could be caused by the robot slipping when contacting the ground, which would also be another reason for the increase of estimation errors for the kinematic localization algorithm. However, the results show a possible absolute trajectory error maximum as the step length increases. Therefore, experiments with longer step lengths are needed to confirm if this maximum exists, if the error tends to stabilize on a given value, or if it keeps increasing indefinitely.

From the results and statistical analysis, as a future work we propose to include the kinematic localization performance curves as cost functions within a footstep planner such as to minimize localization error. This could be used as an active localization algorithm on a robot, in order to optimize its motions to obtain the best localization performance, which could be useful for exploring unknown places, especially those in which there were too few or too many visual textures, and in which visual localization algorithms would not function optimally.

5.7.2 Style and Symmetry

As already presented, changing the style and temporal symmetry of the walking gait increased the kinematic localization (Fig. 5.10). This could be explained by the increase of ground reaction forces on the z axis compared to the normal walking gait, which as we already discussed, are used for the Schmitt trigger to decide when and which foot is in contact with the ground, which in turn is critical for the forward kinematics and therefore for the trajectory estimation.

5.8 Summary

In this chapter we explored how different walking parameters affect the performance of kinematic odometry algorithms, more specifically, the one proposed in [98]. The explored

parameters are step length and walking style and symmetry, and below we present the summary for each tested parameter.

Using the set of experiments with a biped humanoid robot where we changed the step length of the walking patterns, we assessed the effects of this parameter on the performance of a kinematic localization algorithm. Increasing the step length of the walking gaits showed an increase on the acceleration measurements, most likely because faster walking introduced more vibrations on the robot, which affected the performance of the algorithm. This should be cause by the fact that vibrations affect both the readings from the joint encoders as well as from the force/torque sensors on the feet, which in turn would affect the detection of the foot in contact with the ground, introducing error when calculating the forward kinematics of the robot. However, it appears that the error could start decreasing if we increase the step length, for which experiments with longer step lengths will be needed.

From the human point of view, there are studies showing that when humans try to estimate distances using substratal idiothetic cues only, i.e., with the eyes closed, their performance decreases as their walking speed increases. We found a similar behavior with the kinematic odometry algorithm, thus we confirmed with the robot the behavior of distance estimation from substratal idiothetic cues for humans, and hypothesize that this is caused by the increase in vibrations, drift and ground reaction forces forces on the z axis when increasing walking speed.

From the experiments changing the walking style and symmetry of the walking patterns, using a gallop gait and a walking pattern with a different step time for the right and left legs, we found out that kinematic localization was also affected. Both the gallop gait and the asymmetrical gait affected negatively the performance of the kinematic odometry. Ground reaction forces on the vertical axis affected the most, as the kinematic odometry algorithm relies heavily on monitoring the which foot is in contact with the ground, which is then used to calculate the traveled distance from the kinematic chain.

Chapter 6

Conclusion and Future Work

6.1 Contributions of this thesis

In the present thesis, we proposed a perception system based on the human sensory system, focusing on vision, proprioception, and gravitoception, both from the vestibular system on the head, and somatic graviception on the body. The system was then used for the localization of the robot, separating it in the head localization using vision and vestibular inputs, and the body localization using proprioception and somatic graviception. For both system, an EKF approach was used, and particularly for the head localization, a loosely coupled approach was used, together with a sparse visual odometry algorithm, which is best suited for localization purposes. Finally, the effects of different walking parameters on the performance of the localization algorithms from the developed system were assessed experimentally.

We can summarize the contributions of this thesis as follows:

a) Development of a Perception System for a humanoid biped robot, considering gravitational and inertial information from both Vestibular and Somatic systems.

b) Experimental evidence and model for localization performance depending on walking parameters.

   b.1) For Sparse and Direct, Semi-direct and Indirect Visual Odometry Algorithms

   b.2) For a Kinematic Odometry Algorithm

b.3) Assessed walking parameters: step length, walking style and symmetry

c) Possible system to benchmark bipedal robots' localization algorithms

Used two different robots: WABIAN-2R and WALKMAN

6.2 General Discussion

6.2.1 Visual vs Kinematic Localization

In the present thesis, the localization performance of visual and kinematic algorithms was assessed. In general, the results show that kinematic localization is more precise for shorter step lengths, while semi-direct and indirect visual localization perform better for longer step lengths, even taking into account that using a monocular approach introduces the problem of scale ambiguity. Kinematic algorithms suffered much less drift than the visual algorithms, which is an expected outcome given that the sensors used for kinematic localization, i.e. motor encoders and force/torque sensors, have much less uncertainty than the used camera. This point to the fact that not only walking parameters, but also sensing modalities should be taken into account in order to improve the localization performance of the robot. A system able to change the sensor modality depending on the walking parameters or one that combines the information from each sensor modality calibrating the weight of each source of information could be designed to optimize the robot's localization performance. On the other hand, for the scale ambiguity problem of monocular approaches experiments should be made with stereoscopic camera setups, whereas for the drift problem of visual approaches, experiments with higher frame rate cameras a could be useful to obtain a better insight about how to improve their performance.

6.2.2 Robotic Platform Dependence

Although for the validation experiments of the kinematic localization algorithm were made with the humanoid robot WALKMAN, all the experiments to assess the effects of walking parameters on localization were performed with the bipedal humanoid robot WABIAN-2R. Therefore, although the obtained localization performance models are

only true for this robotic platform, performing experiments with other platforms should make it possible to generalize these models for more bipedal humanoid robots.

6.2.3 Humans vs Robots

It is very interesting to have found that changing walking parameters affected the localization performance of the robot, just as it does on humans. Unfortunately, the motions of the robot are still far from those of humans. Specifically, the used step lengths (10 to 25 cm) are too short compared to those humans use. Because of this, a meaningful comparison between humans and the robot was not possible, leaving as an open problem to try closing this gap to enable this comparison. This could shed some light on why humans present differences on the estimation of traveled distances depending on their locomotion parameters, which is still not completely understood.

6.3 Limitations

6.3.1 Experimental setup

One big limitation of this work was the experimental setup, as it was done inside a laboratory. This limited the walking distance for the experiments to 1.5 m on a straight line, which makes it difficult to compare the obtained results with those reported from experiments with humans, where they usually travel longer distances. On the other hand, experiments with humans but with the same conditions as the ones used for the robot could also be performed to enable the comparison of the results.

6.3.2 Robotic platform

Even though the used platform WABIAN-2R was built trying to imitate human proportions, lengths and weights of links, degrees of freedom, range of motion, etc., it still needs improvements to enable faster, wider and more powerful motions. This limited the tested step lengths for the experiments to a maximum of 0.25 m, which for a human would be too small to normally walk. A possible solution is to in the future use a running robot as the one being developed in the Takanishi Laboratory [116], which

should be able to achieve faster and wider motions. Other possible solution is trying to perform the same experiments we did for this study with humans, which would enable the comparison of the results from both humans and the robot.

6.4 Future Work

6.4.1 Human sensory-motor system simulator

Primates, humans included, have a mechanically stabilized visual system, where the image on the retina of the eye is stabilized through reflexes controlled by the vestibular system. Conflicts between the visual and vestibular systems can lead to interesting physiological effects such as motion sickness or vertigo. As the Loosely coupled visuo-inertial approach we used for the head localization system can be considered as virtual stabilized camera with similar functions as a mechanically stabilized visual system, we are planning to explore the possible conflicts and their effects on a robotic system, trying to produce artificial motion sickness or vertigo. Moreover, it could also be possible to simulate effects as the *Mal de Debarquement* [117], which is the persistence of the sense of moving which occurs typically after exposure to such motions, such as after extended travels by train, planes or ships. On the other hand, it is necessary to explore how this sensory-motor system adapts to different situations. For this, approaches using machine learning techniques such as deep learning could be useful in order to train models relating different sensory inputs with motor outputs, depending on the situation, which could be treated as how humans learn to cope with different situations through experience.

6.4.2 Effects of other factors on localization

On this thesis we focused on walking parameters, such as step length, bent or stretched knees, walking style and symmetry. However, there are still other parameters that could have interesting effects, such as walking cadence and step width. Moreover, we are also planning to assess the effects of other motions as walking in non-straight trajectories, hopping, running, etc.

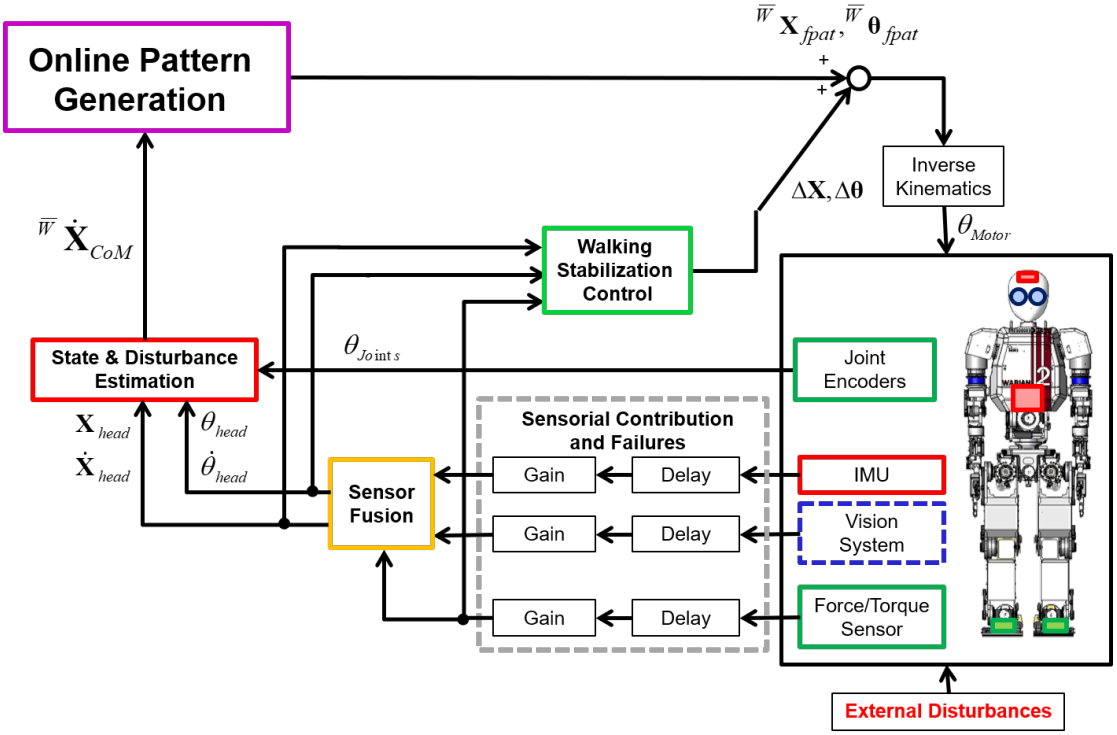


FIGURE 6.1: Proposed system diagram for the Human Sensory-motor System Simulator.

Effects of Walking on Head Localization

Regarding the effects of walking on the visual localization performance of the head localization, we observed a correlation between walking step length and localization performance, as well as between walking style and visual localization performance. From these results, we are planning to include these visual localization performance curves as cost functions within a footstep planner [96] such as to minimize localization error.

About the results obtained from the experiments walking with bent and stretched knees, we are planning to make experiments with more step lengths with stretched knees walking, as the current data does not show any evident difference in visual localization performance in either condition.

Related to the calculation of the scale for monocular VO algorithms, in Chapter 3 we showed that obtaining it from an assumed step length leads to low performance. We are planning to explore ways of extracting the scale before starting the motion and/or during the motion itself, as is the case for some insects that present peering behaviors, or birds that use head-bobbing [118]. Also, another way of solving this problem would be though the use of a stereo camera instead of a monocular one, which we are also

planning for future experiments. Moreover, in this work we focused on sparse visual odometry algorithms and localization, but an interesting extension would be to test dense visual odometry algorithms and see how the mapping performance is affected by walking parameters.

Also, we are interested in exploring how the visual localization performance is influenced by other gait parameters such as stepping time, i.e. the duration of single and double support phases. Other possible parameters to explore are variables that encompass the limbs in coordination such as global phase, which is a variable that seems to be involved in human and perhaps in animal odometry [95].

Effects of Walking on Body Localization

For the effects of walking on the kinematic localization performance of the body localization, we also observed a correlation between walking style and localization performance, and hence we are also planning to include these localization performance curves as cost functions within a footstep planner [96] such as to minimize localization error.

6.4.3 Active localization

Active localization refers to the planning and generation of the motions of a robot prioritizing the localization routine, seeking to increase the efficiency and robustness of localization [119]. Having the experimental models of the localization performance w.r.t. different walking parameters for the robot, it should now be possible to use those models to plan the robot's locomotion in order to minimize the localization error. This could be useful, for example, in tasks where the robot is in an unknown environment, and needs to be as precise as possible to explore the new area, be it for mapping it, or just to traverse it. Also, as the proposed system divides the task into the head and the body, each with different sensors, it could be possible to change the used sensors depending on the situation, e.g. not using vision in the dark, or in very dynamic environments (rain, snow) or with complicated textures (jungle, dessert), relying less on the IMU's when on a moving environment (inside a vehicle) or different gravities (space), etc.

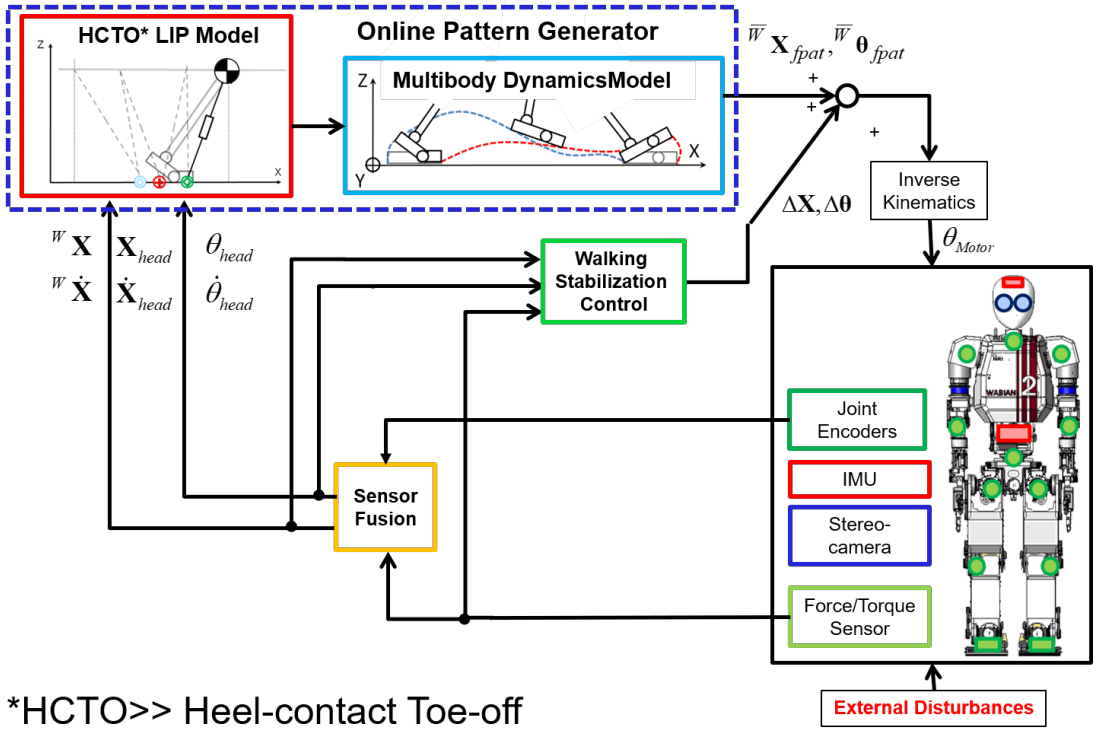


FIGURE 6.2: Robot system overview with the perception system as an input to the heel-contact toe-off gait pattern generator.

6.4.4 Online walking gait pattern generator

On the other hand, we are planning to use the developed localization system as an input to an online waking gait pattern generator. With the system developed in [120], it became possible to regenerate walking patterns to obtain an almost real-time pattern regenerator. Then, as we are interested in humanlike walking parameters, we developed a model considering heel-contact toe-off walking motions to be used in the above mentioned pattern generator. The details are in Appendix D.

References

[1] Yoseph Bar-Cohen. *Biomimetics*. CRC Press, nov 2005. doi: 10.1201/9780849331633.

[2] Julian F.V Vincent, Olga A Bogatyreva, Nikolaj R Bogatyrev, Adrian Bowyer, and Anja-Karina Pahl. Biomimetics: its practice and theory. *Journal of The Royal Society Interface*, 3(9):471–482, aug 2006. doi: 10.1098/rsif.2006.0127.

[3] Bharat Bhushan. Biomimetics: lessons from nature–an overview. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1893):1445–1486, apr 2009. doi: 10.1098/rsta.2009.0011.

[4] Marko Nardini, Dorothy Cowie, A. J. Bremner, D. J. Lewkowicz, and C. Spence. The development of multisensory balance, locomotion, orientation and navigation. *Multisensory Development*, page 137, 2012.

[5] James J. Gibson. *The Ecological Approach to Visual Perception*. Houghton Mifflin, 1979.

[6] OpenStax CNX. Sensory perception - anatomy and physiology - openstax cnx. https://cnx.org/contents/FPtK1zmh@12.8:s3XqfSLV@12/Sensory-Perception, Nov 2018. (Accessed on 01/04/2019).

[7] H Mittelstaedt. Somatic graviception. *Biological psychology*, 42:53–74, 1996.

[8] Todd S. Ellenbecker, George J. Davies, and Jake Bleacher. 24 - proprioception and neuromuscular control. In James R. Andrews, Gary L. Harrelson, and Kevin E. Wilk, editors, *Physical Rehabilitation of the Injured Athlete (Fourth Edition)*, pages 524 – 547. W.B. Saunders, Philadelphia, fourth edition edition, 2012. ISBN 978-1-4377-2411-0. doi: https://doi.org/10.1016/B978-1-4377-2411-0.00024-1.

[9] Barry Wyke. Articular neurology: A review. *Physiotherapy*, 58(3):94–99, March 1972.

[10] J. D. Michelson and C. Hutchins. Mechanoreceptors in human ankle ligaments. *J Bone Joint Surg Br*, 77(2):219–224, Mar 1995.

[11] Kenneth O Johnson. The roles and functions of cutaneous mechanoreceptors. *Current Opinion in Neurobiology*, 11(4):455 – 461, 2001. ISSN 0959-4388. doi: https://doi.org/10.1016/S0959-4388(00)00234-8.

[12] Francis McGlone and David Reilly. The cutaneous sensory system. *Neuroscience & Biobehavioral Reviews*, 34(2):148 – 159, 2010. ISSN 0149-7634. doi: https://doi.org/10.1016/j.neubiorev.2009.08.004. Touch, Temperature, Pain/Itch and Pleasure.

[13] Naoyuki Kakuda and Masanori Nagaoka. Dynamic response of human muscle spindle afferents to stretch during voluntary contraction. *The Journal of Physiology*, 513(2):621–628, dec 1998. doi: 10.1111/j.1469-7793.1998.621bb.x.

[14] R. A. Grunewald, Y. Yoneda, J. M. Shipman, and H. J. Sagar. Idiopathic focal dystonia: a disorder of muscle spindle afferent processing? *Brain*, 120 (Pt 12): 2179–2185, Dec 1997.

[15] M. Hulliger. The mammalian muscle spindle and its central control. *Rev. Physiol. Biochem. Pharmacol.*, 101:1–110, 1984.

[16] I. A. H. Clemens, M. De Vrijer, L. P. J. Selen, J. A. M. Van Gisbergen, and W. P. Medendorp. Multisensory Processing in Spatial Orientation: An Inverse Probabilistic Approach. *Journal of Neuroscience*, 31(14):5365–5377, April 2011. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.6472-10.2011.

[17] R. J. Peterka. Dynamic Regulation of Sensorimotor Integration in Human Postural Control. *Journal of Neurophysiology*, 91(1):410–423, September 2003. ISSN 0022-3077, 1522-1598. doi: 10.1152/jn.00516.2003.

[18] Jack M. Loomis, Roberta L. Klatzky, Reginald G. Golledge, and John W. Philbeck. Human navigation by path integration. In R.G. Golledge, editor, *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*, chapter 5, pages 125–151. 1999. ISBN 9780801859939.

[19] John O'Keefe and Lynn Nadel. *The Hippocampus as a Cognitive Map*. Oxford University Press, 1978. ISBN 0198572069.

[20] Charles R. Gallistel. *The Organization of Learning (Learning, Development, and Conceptual Change)*. A Bradford Book, 1993. ISBN 026257098X.

[21] M. J. Kearns, W. H. Warren, A. P. Duchon, and M. J. Tarr. Path integration from optic flow and body senses in a homing task. *Perception*, 31(3):349–374, 2002.

[22] Timothy M. Ellmore and Bruce L. McNaughton. Human path integration by optic flow. *Spatial Cognition & Computation*, 4(3):255–272, 2004. doi: 10.1207/s15427633scc0403_3.

[23] G. L. Allen, K. C. Kirasic, M. A. Rashotte, and D. B. Haun. Aging and path integration skill: kinesthetic and vestibular contributions to wayfinding. *Percept Psychophys*, 66(1):170–179, Jan 2004.

[24] Gordon Cheng, Sang-Ho Hyon, Jun Morimoto, Ales Ude, Glenn Colvin, Wayco Scroggin, and Stephen C Jacobsen. CB: A Humanoid Research Platform for Exploring NeuroScience. page 6.

[25] A. Ude, V. Wyart, Li-Heng Lin, and G. Cheng. Distributed visual attention on a humanoid robot. In *5th IEEE-RAS International Conference on Humanoid Robots, 2005.*, pages 381–386, Dec 2005. doi: 10.1109/ICHR.2005.1573597.

[26] K. Kaneko, F. Kanehiro, S. Kajita, H. Hirukawa, T. Kawasaki, M. Hirata, K. Akachi, and T. Isozumi. Humanoid robot hrp-2. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, volume 2, pages 1083–1090 Vol.2, April 2004. doi: 10.1109/ROBOT.2004.1307969.

[27] Tran Minh Tuan, Philippe Soueres, Michel Taix, and Emmanuel Guigon. A principled approach to biological motor control for generating humanoid robot reaching movements. In *2008 2nd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics*, pages 783–788, Scottsdale, AZ, USA, October 2008. IEEE. ISBN 978-1-4244-2882-3. doi: 10.1109/BIOROB.2008.4762783.

[28] Tran Minh Tuan, Philippe Soueres, Michel Taix, and Benoit Girard. Eye-centered vs body-centered reaching control: A robotics insight into the neuroscience debate.

In *2009 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 568–573, Guilin, December 2009. IEEE. ISBN 978-1-4244-4774-9. doi: 10.1109/ROBIO.2009.5420609.

[29] Manish N. Sreenivasa, Philippe Soueres, Jean-Paul Laumond, and Alain Berthoz. Steering a humanoid robot by its head. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4451–4456, St. Louis, MO, October 2009. IEEE. ISBN 978-1-4244-3803-7. doi: 10.1109/IROS.2009.5354503.

[30] Shin'ichiro Nakaoka. DARPA Robotics Challenge ni okeru Choreonoid to RT-Middleware no katsuyo, August 2015.

[31] Nobutsuna Endo and Atsuo Takanishi. Development of whole-body emotional expression humanoid robot for ADL-assistive RT services. *Journal of Robotics and Mechatronics*, 23(6):969–977, dec 2011. doi: 10.20965/jrm.2011.p0969.

[32] P. Kryczka, E. Falotico, K. Hashimoto, H. Lim, A. Takanishi, C. Laschi, P. Dario, and A. Berthoz. A robotic implementation of a bio-inspired head motion stabilization model on a humanoid platform. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2076–2081, Oct 2012. doi: 10.1109/IROS.2012.6386177.

[33] T. Pozzo, A. Berthoz, and L. Lefort. Head stabilization during various locomotor tasks in humans. I. Normal subjects. *Exp Brain Res*, 82(1):97–106, 1990.

[34] E. Hirasaki, S. T. Moore, T. Raphan, B. Cohen, and B. Cohen. Effects of walking velocity on vertical head and body movements during locomotion. *Exp Brain Res*, 127(2):117–130, Jul 1999.

[35] T. Imai, S. T. Moore, T. Raphan, B. Cohen, and B. Cohen. Interaction of the body, head, and eyes during walking and turning. *Exp Brain Res*, 136(1):1–18, Jan 2001.

[36] Egidio Falotico, Cecilia Laschi, Paolo Dario, Delphine Bernardin, and Alain Berthoz. Using trunk compensation to model head stabilization during locomotion. In *2011 11th IEEE-RAS International Conference on Humanoid Robots*, pages 440–445, Bled, Slovenia, October 2011. IEEE. ISBN 978-1-61284-868-6 978-1-61284-866-2 978-1-61284-867-9. doi: 10.1109/Humanoids.2011.6100912.

[37] Egidio Falotico, Nino Cauli, Przemyslaw Kryczka, Kenji Hashimoto, Alain Berthoz, Atsuo Takanishi, Paolo Dario, and Cecilia Laschi. Head stabilization in a humanoid robot: models and implementations. *Autonomous Robots*, July 2016. ISSN 0929-5593, 1573-7527. doi: 10.1007/s10514-016-9583-z.

[38] N. G. Tsagarakis, G. Metta, G. Sandini, D. Vernon, R. Beira, F. Becchi, L. Righetti, J. Santos-Victor, A. J. Ijspeert, M. C. Carrozza, and D. G. Caldwell. iCub: the design and realization of an open humanoid platform for cognitive and neuroscience research. *Advanced Robotics*, 21(10):1151–1175, January 2007. ISSN 0169-1864, 1568-5535. doi: 10.1163/156855307781389419.

[39] G. Metta, L. Natale, F. Nori, and G. Sandini. The icub project: An open source platform for research in embodied cognition. In *Advanced Robotics and its Social Impacts*, pages 24–26, Oct 2011. doi: 10.1109/ARSO.2011.6301975.

[40] Elena Franchi, Egidio Falotico, Davide Zambrano, Giovanni Gerardo Muscolo, Laura Marazzato, Paolo Dario, and Cecilia Laschi. A comparison between two bio-inspired adaptive models of Vestibulo-Ocular Reflex (VOR) implemented on the iCub robot. In *Humanoid Robots (Humanoids), 2010 10th IEEE-RAS International Conference on*, pages 251–256. IEEE, 2010.

[41] Alessandro Roncone, Ugo Pattacini, Giorgio Metta, and Lorenzo Natale. Gaze Stabilization for Humanoid Robots: a Comprehensive Framework. *arXiv preprint arXiv:1411.3525*, 2014.

[42] Thomas Mergner, Georg Schweigart, and Luminous Fennell. Vestibular humanoid postural control. *Journal of Physiology-Paris*, 103(3-5):178–194, May 2009. ISSN 09284257. doi: 10.1016/j.jphysparis.2009.08.002.

[43] Karim A. Tahboub. Biologically-inspired humanoid postural control. *Journal of Physiology-Paris*, 103(3-5):195–210, May 2009. ISSN 09284257. doi: 10.1016/j.jphysparis.2009.08.003.

[44] M. Zebenay, V. Lippi, and T. Mergener. Human-like Humanoid Robot Posture Control:. In *Proceedings of the 12th International Conference on Informatics in Control, Automation and Robotics*, pages 304–309, Colmar, Alsace, France, 2015. SCITEPRESS - Science and and Technology Publications. ISBN 978-989-758-122-9 978-989-758-123-6. doi: 10.5220/0005542603040309.

[45] Konrad P. Körding and Daniel M. Wolpert. Bayesian integration in sensorimotor learning. *Nature*, 427(6971):244–247, 2004.

[46] P. Corke, J. Lobo, and J. Dias. An Introduction to Inertial and Visual Sensing. *The International Journal of Robotics Research*, 26(6):519–535, June 2007. ISSN 0278-3649. doi: 10.1177/0278364907079279.

[47] Simon Lynen, Markus W. Achtelik, Stephan Weiss, Margarita Chli, and Roland Siegwart. A robust and modular multi-sensor fusion approach applied to mav navigation. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pages 3923–3929. IEEE, 2013.

[48] Nikolas Trawny and Stergios I. Roumeliotis. Indirect Kalman Filter for 3d Attitude Estimation -A Tutorial for Quaternion Algebra. Technical report, Department of Computer Science & Engineering, University of Minnesota, 2005.

[49] Davide Scaramuzza and Friedrich Fraundorfer. Visual Odometry [Tutorial]. *IEEE Robotics & Automation Magazine*, 18(4):80–92, December 2011. ISSN 1070-9932. doi: 10.1109/MRA.2011.943233.

[50] Friedrich Fraundorfer and Davide Scaramuzza. Visual Odometry : Part II: Matching, Robustness, Optimization, and Applications. *IEEE Robotics & Automation Magazine*, 19(2):78–90, June 2012. ISSN 1070-9932. doi: 10.1109/MRA.2012.2182810.

[51] J. Engel, V. Koltun, and D. Cremers. Direct sparse odometry. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(3):611–625, March 2018. ISSN 0162-8828. doi: 10.1109/TPAMI.2017.2658577.

[52] Different types of visual slam systems — kudan. https://www.kudan.eu/kudan-news/different-types-visual-slam-systems/. (Accessed on 12/26/2018).

[53] Yu Ogura, K. Shimomura, H. Kondo, A. Morishima, T. Okubo, S. Momoki, Hun ok Lim, and A. Takanishi. Human-like walking with knee stretched, heel-contact and toe-off motion by a humanoid robot. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pages 3976–3981, Oct 2006. doi: 10.1109/IROS.2006.281834.

[54] Marc Freese, Surya Singh, Fumio Ozaki, and Nobuto Matsuhira. Virtual robot experimentation platform v-rep: A versatile 3d robot simulator. In Noriaki Ando, Stephen Balakirsky, Thomas Hemker, Monica Reggiani, and Oskar von Stryk, editors, *Simulation, Modeling, and Programming for Autonomous Robots*, pages 51–62, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. ISBN 978-3-642-17319-6.

[55] Morgan Quigley, Ken Conley, Brian P. Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y. Ng. Ros: an open-source robot operating system. In *ICRA Workshop on Open Source Software*, 2009.

[56] C. Forster, Z. Zhang, M. Gassner, M. Werlberger, and D. Scaramuzza. Svo: Semidirect visual odometry for monocular and multicamera systems. *IEEE Transactions on Robotics*, 33(2):249–265, April 2017. ISSN 1552-3098. doi: 10.1109/TRO.2016.2623335.

[57] Inertial measurement unit (imu) devices - lp-research. https://lp-research.com/products/. (Accessed on 12/26/2018).

[58] Dieter Fox, Wolfram Burgard, and Sebastian Thrun. Active markov localization for mobile robots. *Robotics and Autonomous Systems*, 25(3-4):195–207, 1998.

[59] Gabriele Costante, Christian Forster, Jeffrey Delmerico, Paolo Valigi, and Davide Scaramuzza. Perception-aware path planning. *arXiv preprint arXiv:1605.04151*, 2016.

[60] Fara P. Redlick, Michael Jenkin, and Laurence R. Harris. Humans can use optic flow to estimate distance of travel. *Vision research*, 41(2):213–219, 2001.

[61] Jan Faigl. On localization and mapping with RGB-D sensor and hexapod walking robot in rough terrains. In *Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on*, pages 002273–002278. IEEE, 2016.

[62] Martim Brandao, Kenji Hashimoto, José Santos-Victor, and Atsuo Takanishi. Footstep planning for slippery and slanted terrain using human-inspired models. *IEEE Transactions on Robotics*, 32(4):868–879, Aug 2016. ISSN 1552-3098. doi: 10.1109/TRO.2016.2581219.

[63] Olivier Stasse, Andrew J. Davison, Ramzi Sellaouti, and Kazuhito Yokoi. Real-time 3d slam for humanoid robot considering pattern generator information. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 348–355. IEEE, 2006.

[64] SungHwan Ahn, Sukjune Yoon, Seungyong Hyung, Nosan Kwak, and Kyung Shik Roh. On-board odometry estimation for 3d vision-based SLAM of humanoid robot. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 4006–4012. IEEE, 2012.

[65] Liyang Wang, Zhi Liu, C. L. Philip Chen, Yun Zhang, Sukhan Lee, and Xin Chen. A UKF-Based Predictable SVR Learning Controller for Biped Walking. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 43(6):1440–1450, November 2013. ISSN 2168-2216, 2168-2232. doi: 10.1109/TSMC.2013.2242887.

[66] Konstantine Tsotsos, Alberto Pretto, and Stefano Soatto. Visual-inertial ego-motion estimation for humanoid platforms. In *Humanoid Robots (Humanoids), 2012 12th IEEE-RAS International Conference on*, pages 704–711. IEEE, 2012.

[67] Andrew J. Davison and David W. Murray. Mobile robot localisation using active vision. In *Proceedings of the 5th European Conference on Computer Vision - Volume II*, ECCV '98, pages 809–825, London, UK, 1998. Springer-Verlag. ISBN 3-540-64613-2.

[68] Enrique Dunn and Jan-Michael Frahm. Next Best View Planning for Active Model Improvement. In *BMVC*, pages 1–11, 2009.

[69] Sebastian Haner and Anders Heyden. Optimal view path planning for visual SLAM. In *Scandinavian Conference on Image Analysis*, pages 370–380. Springer, 2011.

[70] Sarthak Upadhyay, Ayush Dewan, Arun Kumar Singh, and Madhava Krishna. Trajectory planning for monocular SLAM based exploration system. pages 1–6. ACM Press, 2015. ISBN 978-1-4503-3356-6. doi: 10.1145/2783449.2783476.

[71] Vignesh Prasad, Saurabh Singh, Nahas Pareekutty, Balaraman Ravindran, and Madhava Krishna. SLAM-Safe Planner: Preventing Monocular SLAM Failure using Reinforcement Learning. *arXiv preprint arXiv:1607.07558*, 2016.

[72] R. Sim and N. Roy. Global a-optimal robot exploration in slam. In *2005 IEEE International Conference on Robotics and Automation*, pages 661–666, 2005. doi: 10.1109/ROBOT.2005.1570193.

[73] Christian Mostegel, Andreas Wendel, and Horst Bischof. Active monocular localization: Towards autonomous monocular exploration for multirotor mavs. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 3848–3855. IEEE, 2014.

[74] Ruben Martinez-Cantin, Nando de Freitas, Eric Brochu, José Castellanos, and Arnaud Doucet. A Bayesian exploration-exploitation approach for optimal online sensing and planning with a visually guided mobile robot. *Autonomous Robots*, 27(2):93–103, August 2009. ISSN 0929-5593, 1573-7527. doi: 10.1007/s10514-009-9130-2.

[75] Yifeng Huang and Kamal Gupta. RRT-SLAM for motion planning with motion and map uncertainty for robot exploration. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 1077–1082. IEEE, 2008.

[76] Hiroka Inoue, Masahiro Ono, Sakurako Tamaki, and Shuichi Adachi. Active localization for planetary rovers. In *Aerospace Conference, 2016 IEEE*, pages 1–7. IEEE, 2016.

[77] Teresa A Vidal-Calleja, Alberto Sanfeliu, and Juan Andrade-Cetto. Action Selection for Single-Camera SLAM. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 40(6):1567–1581, December 2010. ISSN 1083-4419, 1941-0492. doi: 10.1109/TSMCB.2010.2043528.

[78] Markus W. Achtelik, Simon Lynen, Stephan Weiss, Margarita Chli, and Roland Siegwart. Motion- and Uncertainty-aware Path Planning for Micro Aerial Vehicles: Motion and Uncertainty Aware Path Planning. *Journal of Field Robotics*, 31(4): 676–698, July 2014. ISSN 15564959. doi: 10.1002/rob.21522.

[79] Seyed Abbas Sadat, Kyle Chutskoff, Damir Jungic, Jens Wawerla, and Richard Vaughan. Feature-rich path planning for robust navigation of MAVs with mono-SLAM. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 3870–3875. IEEE, 2014.

[80] David Gonzalez-Aguirre, Michael Vollert, Tamim Asfour, and Rüdiger Dillmann. Robust real-time 6d active visual localization for humanoid robots. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 2785–2791. IEEE, 2014.

[81] Oliver Lorch, Javier F. Seara, Klaus H. Strobl, Uwe D. Hanebeck, and Günther Schmidt. Perception errors in vision guided walking: analysis, modeling, and filtering. In *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, volume 2, pages 2048–2053. IEEE, 2002.

[82] J.F. Seara and G. Schmidt. Intelligent gaze control for vision-guided humanoid walking: methodological aspects. *Robotics and Autonomous Systems*, 48(4):231 – 248, 2004. ISSN 0921-8890. doi: 10.1016/j.robot.2004.07.003.

[83] Herman van der Kooij, Ron Jacobs, Bart Koopman, and Frans van der Helm. An adaptive model of sensory integration in a dynamic environment applied to human stance control. *Biological cybernetics*, 84(2):103–115, 2001.

[84] Rakie Cham and Mark S Redfern. Changes in gait when anticipating slippery floors. *Gait & Posture*, 15(2):159 – 171, 2002. ISSN 0966-6362. doi: 10.1016/S0966-6362(01)00150-3.

[85] Ann Hallemans, Els Ortibus, Francoise Meire, and Peter Aerts. Low vision affects dynamic stability of gait. *Gait & Posture*, 32(4):547–551, October 2010. ISSN 09666362. doi: 10.1016/j.gaitpost.2010.07.018.

[86] F. Karmali, K. Lim, and D. M. Merfeld. Visual and vestibular perceptual thresholds each demonstrate better precision at specific frequencies and also exhibit optimal integration. *Journal of Neurophysiology*, 111(12):2393–2403, June 2014. ISSN 0022-3077, 1522-1598. doi: 10.1152/jn.00332.2013.

[87] Jonathan Bredin, Yves Kerlirzin, and Isabelle Israël. Path integration: is there a difference between athletes and non-athletes? *Experimental Brain Research*, 167(4):670–674, December 2005. ISSN 0014-4819, 1432-1106. doi: 10.1007/s00221-005-0251-3.

[88] Helen S. Cohen and Haleh Sangi-Haghpeykar. Walking speed and vestibular disorders in a path integration task. *Gait & Posture*, 33(2):211–213, February 2011. ISSN 09666362. doi: 10.1016/j.gaitpost.2010.11.007.

[89] R. Scona, S. Nobili, Y. R. Petillot, and M. Fallon. Direct visual slam fusing proprioception for a humanoid robot. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1419–1426, Sept 2017. doi: 10.1109/IROS.2017.8205943.

[90] Montiel J. M. M. Mur-Artal, Raúl and Juan D. Tardós. ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 31(5): 1147–1163, 2015. doi: 10.1109/TRO.2015.2463671.

[91] S. Choi, J. Park, and W. Yu. Resolving scale ambiguity for monocular visual odometry. In *2013 10th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pages 604–608, Oct 2013. doi: 10.1109/URAI.2013.6677403.

[92] Marie-Luise Mittelstaedt and Horst Mittelstaedt. Idiothetic navigation in humans: estimation of path length. *Experimental Brain Research*, 139(3):318–332, August 2001. ISSN 00144819. doi: 10.1007/s002210100735.

[93] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers. A benchmark for the evaluation of rgb-d slam systems. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 573–580, Oct 2012. doi: 10.1109/IROS.2012.6385773.

[94] Berthold KP Horn. Closed-form solution of absolute orientation using unit quaternions. *JOSA A*, 4(4):629–642, 1987.

[95] M. T. Turvey, C. Romaniak-Gross, R. W. Isenhower, R. Arzamarski, S. Harrison, and C. Carello. Human odometer is gait-symmetry specific. *Proceedings of the Royal Society B: Biological Sciences*, 276(1677):4309–4314, December 2009. ISSN 0962-8452, 1471-2954. doi: 10.1098/rspb.2009.1134.

[96] Martim Brandao, Kenji Hashimoto, José Santos-Victor, and Atsuo Takanishi. Optimizing energy consumption and preventing slips at the footstep planning level. In *15th IEEE-RAS International Conference on Humanoid Robots*, pages 1–7, Nov 2015. doi: 10.1109/HUMANOIDS.2015.7363514.

[97] Adam Bry, Abraham Bachrach, and Nicholas Roy. State estimation for aggressive flight in GPS-denied environments using onboard sensing. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 1–8. IEEE, 2012.

[98] Maurice F. Fallon, Matthew Antone, Nicholas Roy, and Seth Teller. Drift-free humanoid state estimation fusing kinematic, inertial and LIDAR sensing. In *Humanoid Robots (Humanoids), 2014 14th IEEE-RAS International Conference on*, pages 112–119. IEEE, 2014.

[99] N. G. Tsagarakis, D. G. Caldwell, F. Negrello, W. Choi, L. Baccelliere, V.G. Loc, J. Noorden, L. Muratore, A. Margan, A. Cardellino, L. Natale, E. Mingo Hoffman, H. Dallali, N. Kashiri, J. Malzahn, J. Lee, P. Kryczka, D. Kanoulas, M. Garabini, M. Catalano, M. Ferrati, V. Varricchio, L. Pallottino, C. Pavan, A. Bicchi, A. Settimi, A. Rocchi, and A. Ajoudani. WALK-MAN: A High-Performance Humanoid Platform for Realistic Environments: A High-Performance Humanoid Platform for Realistic Environments. *Journal of Field Robotics*, 34(7):1225–1259, October 2017. ISSN 15564959. doi: 10.1002/rob.21702.

[100] R. Smits. KDL: Kinematics and Dynamics Library. http://www.orocos.org/kdl.

[101] Vn-100 - vectornav technologies. https://www.vectornav.com/products/vn-100. (Accessed on 01/05/2019).

[102] T. P. Tomo, S. Somlor, A. Schmitz, S. Hashimoto, S. Sugano, and L. Jamone. Development of a hall-effect based skin sensor. In *2015 IEEE SENSORS*, pages 1–4, Nov 2015. doi: 10.1109/ICSENS.2015.7370435.

[103] P. Mittendorfer, E. Yoshida, and G. Cheng. Realizing whole-body tactile interactions with a self-organizing, multi-modal artificial skin on a humanoid robot. *Advanced Robotics*, 29(1):51–67, 2015. doi: 10.1080/01691864.2014.952493.

[104] B. J. Stephens. State estimation for force-controlled humanoid balance using simple models in the presence of modeling error. In *2011 IEEE International Conference on Robotics and Automation*, pages 3994–3999, May 2011. doi: 10.1109/ICRA.2011.5980358.

[105] Iyad Hashlamon and Kemalettin Erbatur. Center of mass states and disturbance estimation for a walking biped. In *Mechatronics (ICM), 2013 IEEE International Conference on*, pages 248–253. IEEE, 2013.

[106] K. Erbatur, U. Seven, E. Taşkiran, Ö. Koca, M. Yilmaz, M. Ünel, G. Kiziltaş, A. Sabanovic, and A. Onat. Suralp: A new full-body humanoid robot platform.

In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4949–4954, Oct 2009. doi: 10.1109/IROS.2009.5354276.

[107] K. Masuya and T. Sugihara. Dead reckoning of biped robots with estimated contact points based on the minimum velocity criterion. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3637–3642, Nov 2013. doi: 10.1109/IROS.2013.6696875.

[108] K. Masuya and T. Sugihara. Com motion estimation of a humanoid robot based on a fusion of dynamics and kinematics information. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3975–3980, Sep. 2015. doi: 10.1109/IROS.2015.7353937.

[109] S. Nakaoka, S. Hattori, F. Kanehiro, S. Kajita, and H. Hirukawa. Constraint-based dynamics simulator for humanoid robots with shock absorbing mechanisms. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3641–3647, Oct 2007. doi: 10.1109/IROS.2007.4399415.

[110] X. Xinjilefu, Siyuan Feng, Weiwei Huang, and Christopher G. Atkeson. Decoupled state estimation for humanoids using full-body dynamics. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 195–201. IEEE, 2014.

[111] Nicholas Rotella, Michael Bloesch, Ludovic Righetti, and Stefan Schaal. State estimation for a humanoid robot. In *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pages 952–958. IEEE, 2014.

[112] Michael Bloesch, Marco Hutter, Mark A. Hoepflinger, Stefan Leutenegger, Christian Gehring, C. David Remy, and Roland Siegwart. State estimation for legged robots-consistent fusion of leg kinematics and IMU. *Robotics*, page 17, 2013.

[113] Jorhabib Eljaik, Naveen Kuppuswamy, and Francesco Nori. Multimodal sensor fusion for foot state estimation in bipedal robots using the Extended Kalman Filter. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 2698–2704. IEEE, 2015.

[114] S. Lohmeier, T. Buschmann, and H. Ulbrich. System design and control of anthropomorphic walking robot lola. *IEEE/ASME Transactions on Mechatronics*, 14(6):658–666, Dec 2009. ISSN 1083-4435. doi: 10.1109/TMECH.2009.2032079.

[115] R. Wittmann, A. Hildebrandt, D. Wahrmann, D. Rixen, and T. Buschmann. State estimation for biped robots using multibody dynamics. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2166–2172, Sep. 2015. doi: 10.1109/IROS.2015.7353667.

[116] T. Otani, K. Hashimoto, M. Yahara, S. Miyamae, T. Isomichi, M. Sakaguchi, Y. Kawakami, H. O. Lim, and A. Takanishi. Running with lower-body robot that mimics joint stiffness of humans. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3969–3974, Sep. 2015. doi: 10.1109/IROS.2015.7353936.

[117] Timothy C. Hain, Philip A. Hanna, and Mary A. Rheinberger. Mal de debarquement. *JAMA Otolaryngology–Head & Neck Surgery*, 125(6):615–620, 06 1999. ISSN 2168-6181. doi: 10.1001/archotol.125.6.615.

[118] Alfred Bruckstein, Robert J. Holt, Igor Katsman, and Ehud Rivlin. Head movements for depth perception: Praying mantis versus pigeon. *Autonomous Robots*, 18(1):21–42, 2005.

[119] Wolfram Burgard, Dieter Fox, and Sebastian Thrun. Active mobile robot localization. In *Proceedings of the Fifteenth International Joint Conference on Artifical Intelligence - Volume 2*, IJCAI'97, pages 1346–1352, San Francisco, CA, USA, 1997. Morgan Kaufmann Publishers Inc. ISBN 1-555860-480-4.

[120] Przemyslaw Kryczka, Yukitoshi Minami Shiguematsu, Petar Kormushev, Kenji Hashimoto, Hun-ok Lim, and Atsuo Takanishi. Towards dynamically consistent real-time gait pattern generation for full-size humanoid robots. In *Robotics and Biomimetics (ROBIO), 2013 IEEE International Conference on*, pages 1408–1413. IEEE, 2013.

[121] P.E. Klopsteg and P.D. Wilson. *Human Limbs and Their Substitutes*. Hafner Press, 1968. ISBN 9780028479507.

[122] Przemyslaw Kryczka, Yukitoshi Minami Shiguematsu, Petar Kormushev, Kenji Hashimoto, Hun-ok Lim, and Atsuo Takanishi. Towards dynamically consistent real-time gait pattern generation for full-size humanoid robots. In *Robotics and Biomimetics (ROBIO), 2013 IEEE International Conference on*, pages 1408–1413, Dec 2013. doi: 10.1109/ROBIO.2013.6739663.

[123] A. Takanishi, Hun ok Lim, M. Tsuda, and I. Kato. Realization of dynamic biped walking stabilized by trunk motion on a sagittally uneven surface. In *Intelligent Robots and Systems '90. 'Towards a New Frontier of Applications', Proceedings. IROS '90. IEEE International Workshop on*, pages 323–330 vol.1, Jul 1990. doi: 10.1109/IROS.1990.262408.

[124] Shuuji Kajita and Kazuo Tani. Study of dynamic biped locomotion on rugged terrain-derivation and application of the linear inverted pendulum mode. In *Proc. ICRA Robotics and Automation IEEE Int. Conf*, pages 1405–1411, 1991. doi: 10.1109/ROBOT.1991.131811.

[125] S. Kajita, O. Matsumoto, and M. Saigo. Real-time 3d walking pattern generation for a biped robot with telescopic legs. In *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, volume 3, pages 2299–2306 vol.3, 2001. doi: 10.1109/ROBOT.2001.932965.

[126] Jacqueline Perry. *Gait Analysis: Normal and Pathological Function.* SLACK, Thorofare, NJ, 1st edition, 1992.

Appendix A

Embodiment Informatics

Let us start from the beginning. A quick search in a dictionary can give us the meaning of each word. First let us see "embodiment":

> Embodiment (n.): someone or something that is a perfect representative or example of a quality, idea, etc.
> : the act of embodying: the state of being embodied. [1]

Which leads us to look for "embody":

> Embody (v.): to represent (something) in a clear and obvious way: to be a symbol or example of (something)
> : to include (something) as a part or feature. [2]

And then "informatics":

> Informatics (Information science) (n.): Discipline that deals with the processes of storing and transferring information. It attempts to bring together concepts and methods from such varied disciplines as library science, computer science and engineering, linguistics, and psychology to develop techniques and devices to aid in the handling of information... [3]

[1] "embodiment." Merriam-Webster.com. 2011. http://www.merriam-webster.com (31 July 2014).
[2] "embody." Merriam-Webster.com. 2011. http://www.merriam-webster.com (31 July 2014).
[3] "informatics." Merriam-Webster.com. 2011. http://www.merriam-webster.com (31 July 2014).

The first notorious thing looking at the definition of "informatics" is that it is a very broad field of study, which has been growing since its appearance, and has been greatly improved along the development of new, better and faster technologies, now covering almost every aspect of our lives; information is the core of all existing knowledge and technologies. We can then analyze "embodiment" and try to join the definitions: embodiment informatics should be the perfect representative, clear and obvious, of the processing, storing and transferring of information. But this definition seems weak, as the amount of information nowadays is very big, so I would also include the fact that embodiment informatics should deal with how to use all this information for the creation of new useful technologies for us. But even with that, the concept continues being somehow ambiguous and vague, as informatics reach is too wide, and so the possible embodiments. So, after having examined the definitions and tried to come up with a possible definition for embodiment informatics, let us put it into context. As mentioned above, today we live in a world where information about anything and everything can be obtained, and where with the right tools, this information can be used to come up and create a vast variety of objects. Then, an efficient way of obtaining and processing the right information, as well as an effective way to choose the proper tools becomes necessary. For the first part of obtaining and processing the information, the informatics comes into play, as it searches for better, faster and more robust techniques to acquire, process and manipulate information. As for the second part of choosing the proper tools, I think that is where the field of expertise of each one of us comes into play. So, as I can only speak for my field of research, I will now move on to my own meaning of embodiment informatics. As a mechatronics engineer, one must have strong foundations on precision mechanics, electronic circuits and control software, to ultimately be able to combine these to build more complex systems. One of the most common applications is robotics, and that is exactly where I am. Robotics is still very wide, but in general it comprises a sensing system, to gather required information from the environment, the control system, which processes this gathered information and transforms it into orders for the actuation system, which is the one that accomplishes the objectives of the overall system. My interest is on humanoid robots, i.e., robots that can mimic humans from various perspectives, and I am currently working on walking biped humanoid robots. These are complex systems, as the action of biped walking is a difficult motion which requires to be rightly stabilized to achieve it. Since the 1960's, there had been researches to realize this, and today we have many examples of successful bipedal

walking humanoid robots. The research that led to the current state, also helped the better understanding of our own, which I think is very important, but there is still a lot of space for improvement. As a future vision, I would like to see humanoid robots helping us in various task on our daily lives, while moving seamlessly through any kind of environments. To achieve this, it is necessary for them to efficiently gather all the possible information of the environment using many kinds of sensors, to then combine all this sensory information and process it to finally generate adequate commands for its end effectors. But the information that can possibly be obtained from the environment is a lot, and having that much information means that the time to process it will also be a lot. That is where informatics could help to maximize the amount of processable information to get the best possible commands for the actuators. On the other hand, with the appearance of these kind of robots, it will be possible to interact with them which means more information to process, but also new possible areas to create new devices. One idea is to have many kinds of wearable sensors for humans, that could somehow be read by different systems, to ultimately generate completely new ways of human-robot interactions. And in the end, if this could be converted into a virtuous cycle, where with new technologies, better robots could be produced, and with these, new ways of cooperation and interaction, and so on, that I think would be the ideal case. To conclude, my definition of embodiment informatics is the synergic cooperation between robotics sciences and informatics sciences, in order to combine the best of both to generate innovation in many fields and finally bring a better quality of life to humans.

Appendix B

Biped Robots in Waseda

The University of Waseda was the pioneer of research on the field of humanoid robotics. All started in 1966, and the next year, in 1967, the first of a series of biped robots called Waseda Legs (WL) came out. WL-1 was built on the basis of the analysis of lower limbs locomotion, and it would be a milestone for this field of research. In 1969, the WL-3, a new model of lower limbs actuated by electro-hydraulic servos which achieved human-like movements was developed. Later came the WL-5, able to bend its body laterally, and to move its center of gravity on the frontal plane, through the use of eleven mechanical degrees of freedom (DoF), five in each leg and one on the trunk. This model accomplished automatic biped walking and direction changing, which lead it to be used as the lower limbs of the world's first full-scale anthropomorphic robot, WABOT-1 (WAseda roBOT-1), which was developed by Ichiro Kato and co-workers in 1973. WABOT-1 (Fig. B.1) was hydraulically powered, and had the ability to measure distances of objects thanks to external receptors similar to eyes and ears, could communicate in Japanese and was able to grip and transport objects with tactile-sensor mounted hands. However, its disproportionately large feet made it shuffle, rather than walk, realizing "static walking" (45 s/step).

Then, in 1979, WL-9 achieved quasi-dynamic walking (10 s/step) for the first time in the world. It used a 16-bit microcomputer instead of the frequently used minicomputers. In 1983, with carbon fiber reinforced plastic (CFRP) and introducing the use of rotary servo-actuators, the WL-10R (for Refined) managed plane walking, which consists on walking laterally, forward, backward and turning thanks to an added DoF in the yaw

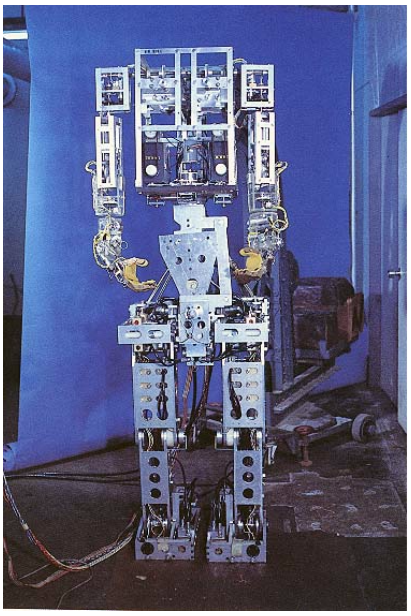


FIGURE B.1: WABOT-1 (1973)

FIGURE B.2: WABOT-2 (1984)

axis of the hip joint, and reached a walking speed of 4.4 s/step. A year later, in 1984, torque sensors would be attached to the ankle and hip joints of WL-10RD, allowing torque-feedback control, and realizing once again, for the first time, complete dynamic walking (1.3 s/step). Also, dynamic walking with a step time of 2.5 s/step on stairs and inclined surfaces was achieved. On that same year, the second of the full-scale robot series, WABOT-2 was developed (Fig. B.2). This robot was defined as a "specialist robot", in contrast to WABOT-1 which was described as a "versatile robot", as it was aimed to accomplish a particular goal: to be a a musician humanoid robot. WABOT-2 was able to play the keyboard by reading musical sheets, and also as to accompany while listening to a singing person. This robot was a milestone in the development of a "personal robot". In 1986, going back to the legs series, came WL-12, a hydraulically actuated biped robot, equipped with a two-DoF waist and an upper body, designed to compensate for the moments generated by the lower limbs, achieving a more human-like motion with improved walking stability. In 1993 the WL-12RVII (Fig. B.3), an upgraded version of the WL-12 was developed that introduced a new adaptive control method, was able to maintain stable dynamic walking in unknown paths and stairs in a human residential environment. The WL series continued until 2007, with the last WL-16RV (Fig. B.5), which now used parallel mechanism based legs, was able to carry and transport a person of up to 75 kg and could cope with external disturbances of up to 105 N.

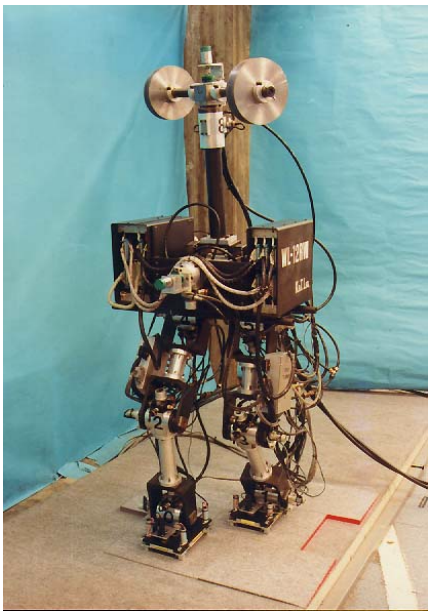


FIGURE B.3: WL-12RVII (1993)

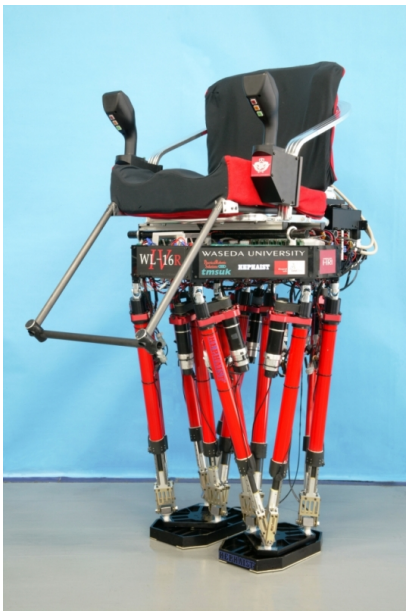


FIGURE B.4: WL-16RIV (2006)

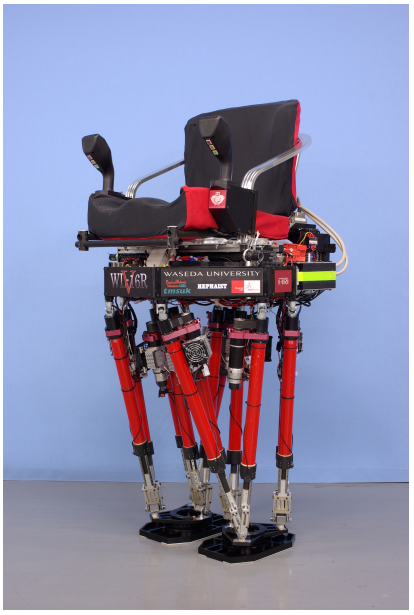


FIGURE B.5: WL-16RV (2007)

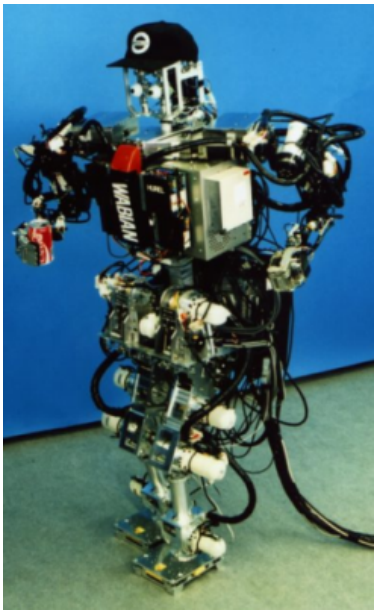


FIGURE B.6: WABIAN (1996)

On the other hand, in 1996 the WABIAN (WAseda BIpedal humANoid) series started, developed under the following criteria:

1. The size of the robot should be that of an average Japanese adult woman, to make collaborative work with humans possible.
2. The walking speed should be approximately that of humans.
3. The robot should have a 3 DoF trunk and 6 DoF arms.

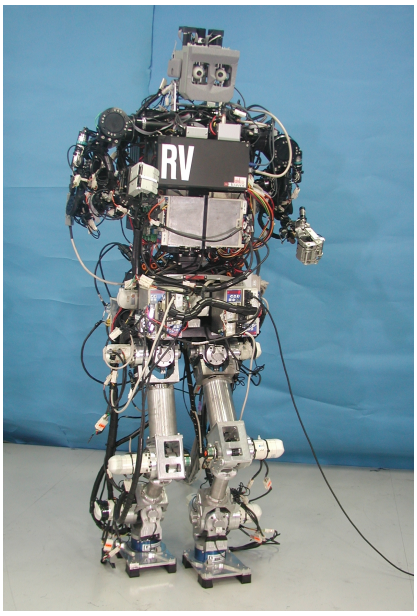


FIGURE B.7: WABIAN-RV (2001)

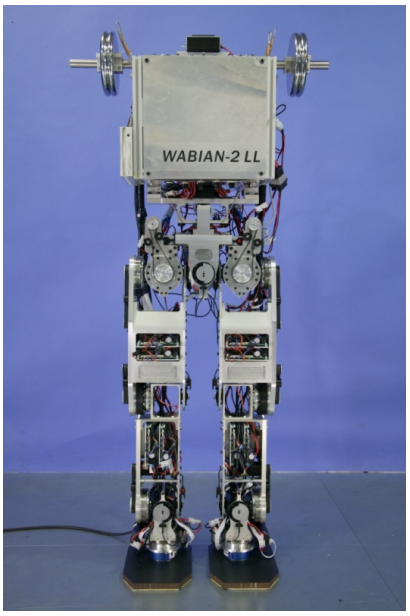


FIGURE B.8: WABIAN-2/LL (2003)

4. The joints should be actuated by servomotors.

5. The control computer and motor drives should be installed on-board, excluding the power supply.

The first WABIAN (Fig. B.6) had 35 mechanical DoF in total: two three DoF legs, two ten DoF arms, a two DoF neck, four DoF in the eyes and a three DoF waist, and was capable of doing various walking sequences and dancing, it could be operated at a distance via Integrated Services Digital Network (ISDN) lines etc.

WABIAN was upgraded, adding mechanical DoF, making it able to generate walking patterns online, enabling its control through a voice recognition system, etc. until 2001, when WABIAN-RV (Fig. B.7) was developed.

In 2004, an upper body with upper limbs were attached to WABIAN-2 (Fig. B.9), which made it capable of simulating various human motions. With this new capability, WABIAN-2 has been refined (Fig. B.10), and its researches have focused on studying the human locomotion in order to make the robot's motions more human-like, achieving heel-contact toe-off walking, evaluation of a walking assistance device, simulation of a walking gait of an hemiplegic patient, walking with feet mechanisms which modeled the human feet arches, turning on one foot's heel and the other's toe using a slipping motion, among others. Also, there are researches on walking stabilization on various surfaces

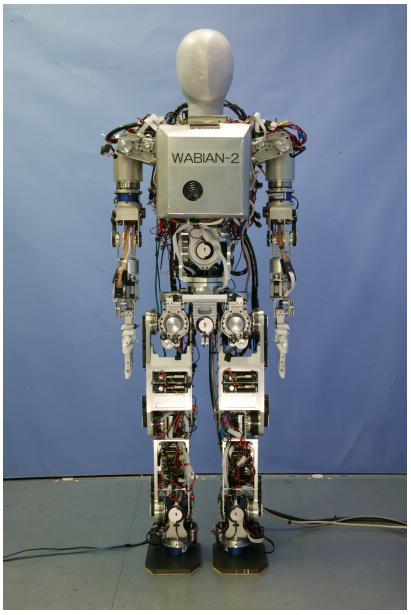


FIGURE B.9: WABIAN-2 (2004)

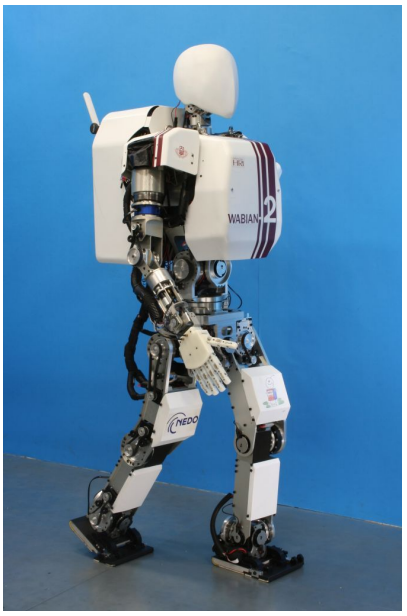


FIGURE B.10: WABIAN-2R (2005)

as slopes, uneven terrains and soft grounds, as well as works on online and real time walking pattern generation.

TABLE B.1: History of Waseda Leg Series 1/2

Year	Robot Name	Contents and Effects of Research
1967	WL-1	Analysis and modeling of human lower limbs basic data
1968	WL-2	Analysis of human joint angles at level ground walking
1969	WL-3	Realization of standing and sitting position with master/slave system
1970	WL-4	Upper body leveling with a 35 kg load
1971	WL-5	Implementation of manual rectilinear walking and direction changing
1972	WL-5	Realization of "static walking"
1973	WL-6D	Proposal of ZMP control
1974	WL-7D	Continuation of ZMP control
1975-7	WL-8D	Continuation of ZMP control
19678	WL-8D	Proposal of quasi-dynamic walking control method
19679	WL-9D	
1980	WL-9DR	Realization of quasi-dynamic walking
1981	WL-9DRII	Realization of quasi-dynamic walking
1982	WL-10	
1983	WL-10R	Realization of quasi-dynamic plane walking
19684	WL-10RD	Realization of parallel quasi-dynamic walking
1983-85	WHL-11	Designed by Kato laboratory, built by Hitachi, Ltd., realization of "static walking" and exhibition in international scientific exposition
19685	WL-10RD	Dynamic stairs climbing (0.1 m) and descent (0.05 m)
19686	WL-12	Implementation of dynamic walking with upper body motion compensation
1987	WL-12R	Dynamic walking under well-known trapezoidal external force of 100 [N]
1988	WL-12RII	Dynamic stepping on-site under unknown external forces of up to 100 [N] during 0.16 [s]
1989	WL-12RIII	Implementation of a real-time compensation for a dynamic stepping on-site under unknown external forces of up to 100 [N] for 0.3 [s] Dynamic walking on slopes of ± 10 [deg] and climbing and descent of stairs of up to 0,1 [m]
1990	WL-12RIV	Dynamic stepping on-site under unknown external forces from the sides of up to 100 [N] during 0.16 [s] Proposal of an upper body trajectory learning control system based on actual ZMP measurement
1991	WL-12RV	Dynamic walking under unknown external forces of up to 100 [N] during 0.18 [s] Realization of a high-speed dynamic walking (0.3[m] of step length and 0.54[s/step])
1992	WL-12RVI	Walking on horizontal uneven surfaces (max. 12 [mm] high) Proposal of lower limbs optimal trajectory planning algorithm and trajectory learning control system
	WL-12RH	Proposal of an upper body trajectory learning walking control system based on actual ZMP measurements and human assistance force information

TABLE B.2: History of Waseda Leg Series 2/2

Year	Robot Name	Contents and Effects of Research
1993	WL-12RVII	Realization of walking on a flat surface with trapezoidal obstacles with ±3 [deg] of inclination
	WL-12RT	Realization of 0.15 [m] step length 2.5 [s/step] walking, with dynamic turning motions of up to 45 [deg] in 2 [s]
1994-5	WL-13	Antagonistically driven joint bipedal quasi-dynamic walking with a step length of 0.10 [m] and 7.68 [s/step]
1996	WL-14	Antagonistically driven joint bipedal dynamic walking
1997		Realization of antagonistically driven joint bipedal dynamic walking based on human gait model
2001	WL-15	Adoption of a generalized Stewart platform parallel mechanism for the legs Realization of dynamic walking (step length of 0.2 [m], 0.80 [s/step]), high-speed dynamic turning motions (up to 90 [deg] in 2.0 [s]), dynamic walking with an added load of 18 [kg] and attitude maintenance without energization.
2002	WL-16	Walking on uneven surfaces using virtual compliance control with a step lenght of 0.3 [m] Walking with an added load of 50 [kg] and a step length of 0.1 [m]
2003	WL-16R	First robot to dynamically walk carrying an adult person (male approx. 60 [kg], female approx. 50 [kg Realization of dynamic walking with an approx. 100[kg] passenger using an own-weight support torque reduction mechanism Walking on surfaces with unevennesses of 8 [mm] and a slope of 3 [deg]
2004	WL-16RII	Climbing and descending stairs carrying an adult person Stable walking on surfaces with ±20[mm] unevennesses by using a landing trajectory correction control
2005	WL-16RIII	Introduction of a walking pattern generator with a passenger model
2006	WL-16RIV	Increase of the possible added load to 75 [kg] by optimizing the links localization and reducing the weight of the robot Straight walking under unknown frontal perturbations of up to 105[N] by introducing a disturbance compensation control
2007	WL-16RV	Foot landing impact reduced to half by the use of the landing impact relaxation, uneven terrain adaptation control Walked with an added load of 75 [kg] for 75 [min] using a cooling mechanism

TABLE B.3: History of Waseda Humanoid Robot Series 1/2

Year	Robot Name	Contents and Effects of Research
1972	WABOT-1	First intelligent humanoid robot, able to make simple conversations in Japanese, bipedal walking and objects grasping
1984	WABOT-2	Keyboardist humanoid robot, able to make natural conversations in Japanese, music sheet recognition using vision and also able to perform accompanying a singing human by sound recognition
1996	WABIAN	Biped humanoid robot with two arms, a head and a audio-visual information acquisition system, able to walk on flat surfaces, carry objects with its arms and dancing. Able to exchange gestures with people, and communicate remotely via ISDN lines.
1997		Introduction of a six axis force/torque sensing system, making it able to walk with a step length of 0.1 [m] ans 1.28 [s/step] continuously depending on forces exerted on it hands.
	WABIAN-R	Improvement of the legs to 6 DoF and realization of bipedal walking using waist and trunk.
1998	WABIAN-RII	Achievement of expressive motions through the use of redundant DoF on the entire body.
1999	WABIAN-RIII	Walking stabilization introducing a compliance control in the ankle joints.
2000	WABIAN-RIV	Development of a real time pattern generation system that relies on visual and auditory sensory inputs. Static disturbance compensation.
2001	WABIAN-RV	Development of auditory information acquisition system, and application of this to the real time pattern generation system.
2002	WABIAN-2 LL	Development of lower limbs with 7 DoF per leg and a 2 DoF waist. Development of an extended knees walking pattern generator.
2003	WABIAN-2/LL	Improvement of mechanical stiffness.
2004	WABIAN-2	Addition of a 2 DoF trunk and 7 DoF arms. Introduction of limb compliance control. Experiments using a walking assistance device, proving the validity of the robot as a human motion simulator.
2005	WABIAN-2R	Realization of heel-contact toe-off walking using a passive toe joint foot mechanism.
2006		Introduction of a battery charge/discharge circuit. Development of an algorithm to mimic an hemiplegic person's walking, based on the analysis of the data from a patient.

TABLE B.4: History of Waseda Humanoid Robot Series 2/2

Year	Robot Name	Contents and Effects of Research
2007	WABIAN-2R	Walking on-site on a 4[deg] slope using a foot mechanism to cope with uneven terrains. Constant hip height walking over 20[mm] unevennesses inside the laboratory, and a 4[deg] slope outside, using an attitude control based on a gyro sensor. Walking over unevennesses of 15[mm] in the roll direction and 20[mm] 15[mm] in the pitch direction, with the combination of a surface information detection foot mechanism and attitude control.

Appendix C

Robotic Platform WABIAN-2R

C.1 Basic Design

This robot was initially designed to be used as a human motion simulator, and to be consistent with that, its dimensions were also chosen to coincide with those of a human individual. In the case of WABIAN-2R, it was constructed based on the dimensions of an average Japanese adult woman as can be seen in Fig. C.12, having 1500 mm of height, and 60 kg of weight.

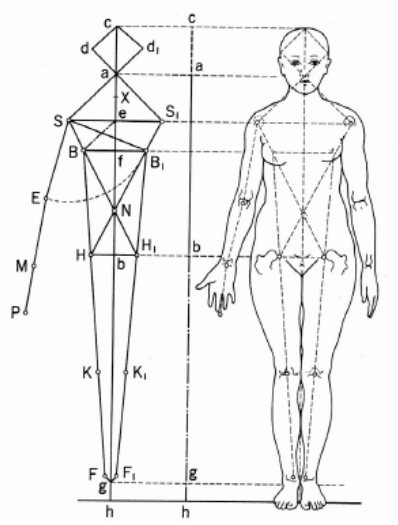


FIGURE C.1: Dimensions of an average Japanese adult woman.

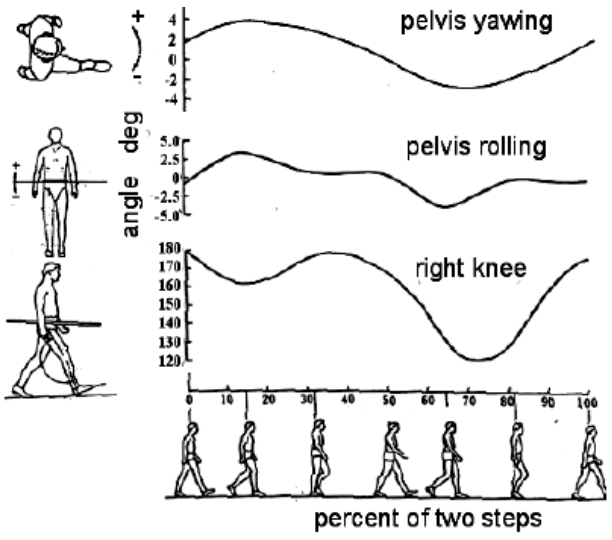


FIGURE C.2: Pelvis and knee motion during walking, from [121].

As the design objective of the robot was to be a human motion simulator, an analysis of the human motion was necessary to then transfer those to the robot. Humans move thanks to the musculoskeletal system, also known as locomotor system, which is composed of bones that act as rigid links, cartilages that join bones and form joints, and finally muscles and tendons, the first used to actuate the different body parts generating motion, and the latter used to join and transfer the forces from muscles to bones. This system is very complex, and although progress ha been made to develop artificial systems imitating the human musculoskeletal system mechanisms, it is still far from being actually useful for robotic motion. Therefore, the primary goal of the mechanical design was determined to be the development of a robot capable of accomplishing motions equivalent to human motions, i.e., capable of imitating human motions, giving special attention to walking.

Klopsteg et al. [121] proposed the analysis of human normal gait, for which they used healthy people without any physical or mental handicaps as subjects. They observed pelvis motions in the frontal plane (which we will call roll motion in the present study) and the horizontal plane (which we will call yaw motion), but did not observe much motions of the waist in the side plane (which we will call pitch motion). Also, they realized that humans can walk with stretched knees while maintaining the height of upper body constant (Fig. C.2). Therefore, a humanoid robot capable of imitating human walking should be able to move its pelvis in the roll and yaw axes, maintaining these motions independent from the trunk.

Another study of gait analysis and biomechanics showed that in steady walking, the pubic symphysis allows a crank-like motion for the pelvis. Based on this, it was decided to make the robot able to also achieve this motion while walking (Fig. C.3).

From all the above, as the present humanoid robot should perform as close as possible to an actual human being, for the walking motion it should be able to move its hip in roll and yaw axes, showing a crank-like motion for its pelvis, and not only that, but it also needs these hip motions to be independent from the robot's trunk movements. The structure of the human pelvis served as a model for the design of this robot's waist.

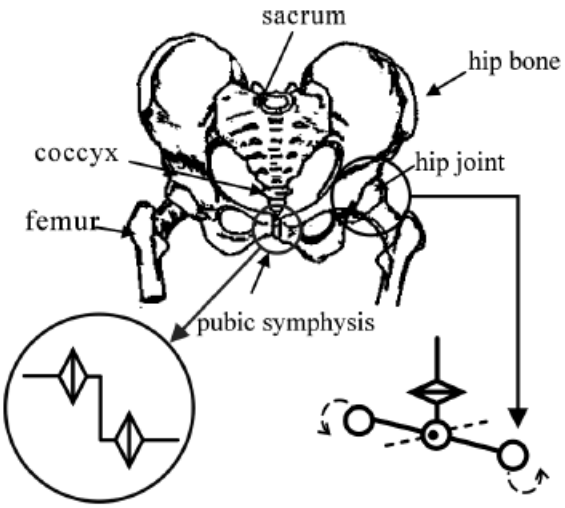


FIGURE C.3: Human pelvis, and the crank-like motion model.

C.2 Degrees of Freedom and Configuration

In Fig. C.4, the configuration of the robot's degrees of freedom is presented. The reference coordinate system is fixed to the ground, and the directions are as presented in the same figure.

For the limbs, the robot has two 6 DoF legs and two 7 DoF arms. For the waist and trunk, it has a 2DoF waist where the roll axis and the yaw axis are perpendicular to each other and cross in the middle point between the hip joints, and the yaw joint is laid on the trunk side, so that it can be used either to rotate the waist or the trunk. As for the trunk, it has 2 DoF, but as mentioned, can make use of the waist yaw motion too. Finally, it has a 3 DoF neck, making a total of 33 DoF.

C.3 Mechanisms

To achieve the best performance while on motion, the robot's mechanical components should be lightweight, highly stiff and should allow wide ranges of motion. For this, the frames were build from duraluminium. The actuator systems were designed to achieve compact human-like joint mechanisms, for which we used DC motors, strain wave gearings, timing belts and pulleys. This accomplished a double speed reduction mechanism, which removed the need to align the motor and the joint axes, and allowed high reduction ratios. Also, the high reduction ratio of the strain wave gearings and

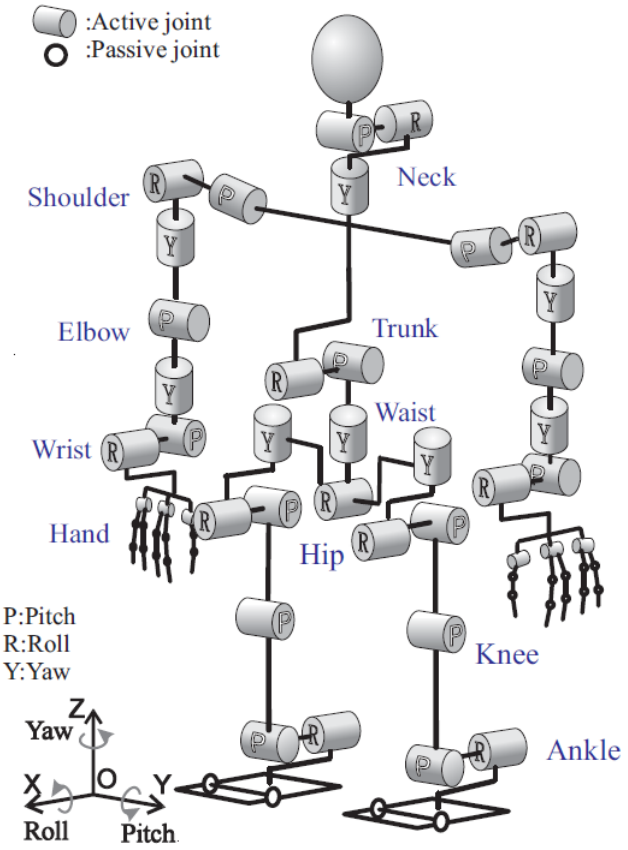


FIGURE C.4: Degrees of freedom configuration of WABIAN-2R.

their placement just before the end effectors helped to reduce the errors coming from the motors and those caused by the belts elasticity, while their no-backdrivability prevented the motors from receiving damage from perturbations on the end effectors. The general diagram of actuation systems os shown in Fig. C.5.

Software simulations based on the Newton-Euler method and the mass distribution of the robot were carried out to determine the specifications of each joint, as maximum torque and maximum rotating speed. These will be detailed in the following sections.

The lower limbs are the basic structure of this robot, and are designed to resist the load of the upper body.

To design the middle and upper section of the body, i.e. waist and trunk, first they were modeled in the sagittal and frontal plane. Fig. C.6 shows the model used for the trunk. Based on these models, two kinds of simulations were done. First the models were considered in a static situation to determine the maximum angles. Then, considering the dynamic case for the models, the angle progression during a walking cycle was

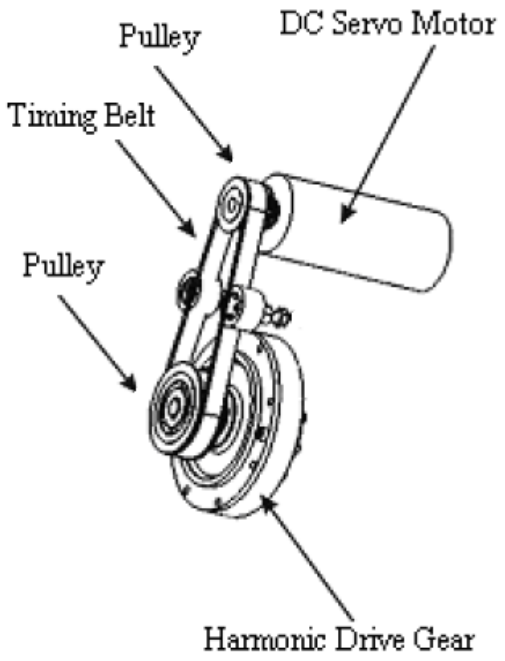


FIGURE C.5: Actuation system used in the robot.

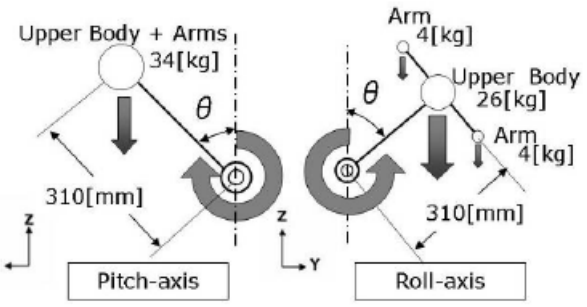


FIGURE C.6: Static models of the trunk in sagittal (left) and frontal (right) planes.

determined. This second type of simulation was carried out based on the fact that while walking, the trunk should move in such a way to keep the balance of the upper body.

The upper limbs of the robot have 7 DoF each, and were designed firstly to support the robot's balance while walking, and secondly, to be able to use the robot to evaluate walking assist devices. For the second goal, the design objective for the arms was to be able to hold almost all of the robot's weight while leaning on the elbows on one of such devices.

All the mechanical design was done using the 3D CAD software, SolidWorks®.

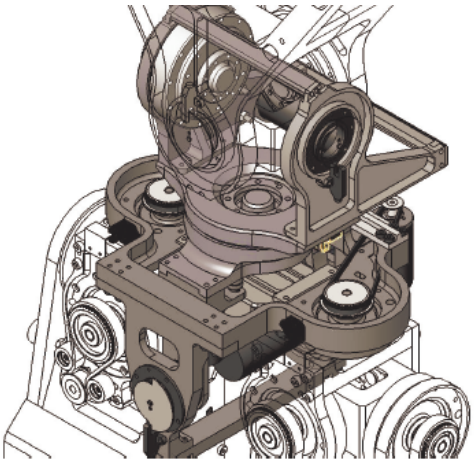


FIGURE C.7: Waist mechanism of WABIAN-2.

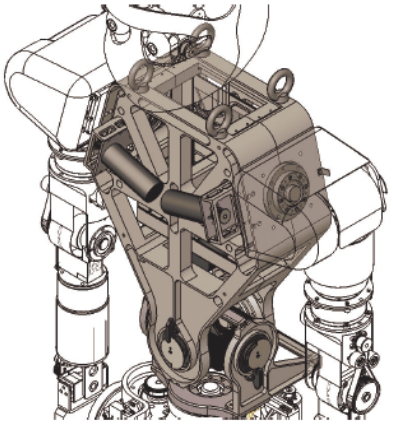


FIGURE C.8: Trunk mechanism of WABIAN-2.

C.4 Control System

The control of the robot is performed by a control system programmed in a computer mounted on the robot's back, resembling a backpack. This computer consists of a half-sized CPU board and data acquisition boards, interconnected by a backplane PCI board. The data acquisition boards are HRP interface boards (Table C.1) and NITTA F/T (force/torque) sensor interface boards. The OS installed in the control computer is a Linux based real-time operating system called QNX. The general control system of WABIAN-2R can be seen in Fig. C.9.

The sensing system of the robot comprises incremental encoders mounted to the shaft of each motor, photointerruptors to initialize the angular position of the motor encoders, and six axis force/torque sensors (Fig. C.11, Table C.2) on the feet, mounted between the ankle joint and foot plate, and on the arms, between the elbow joint and the upper arm.

As for the control of the DC motors of the actuation systems, servo-motor drivers developed by TOKUSHU DENSO Co., Ltd. were used. The models used are the TD12770-48W10 and the TD12770-48W0, which have a maximum output power of 1200 W and 600 W respectively, with an input of 48 V. They use Metal-Oxide Semiconductor Field-Effect Transistor (MOSFET) technology, and enable speed and torque control, not requiring a tachometer generator as it has an in-built electrical governor (speed limiter) for the speed control. It can generate a PWM signal up to 100 kHz, and they can be observed in Fig. C.10.

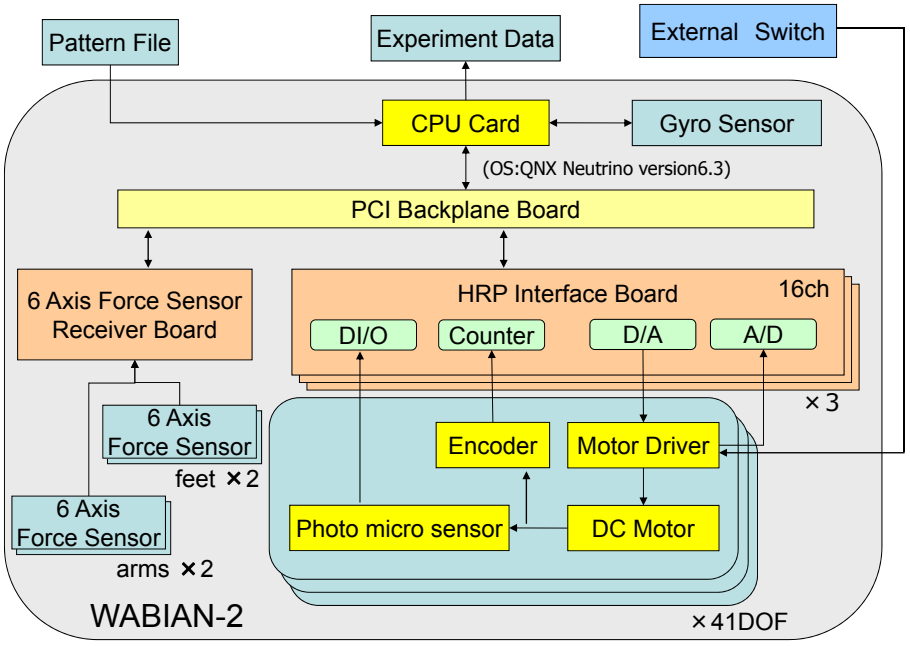


FIGURE C.9: General control system of WABIAN-2R.

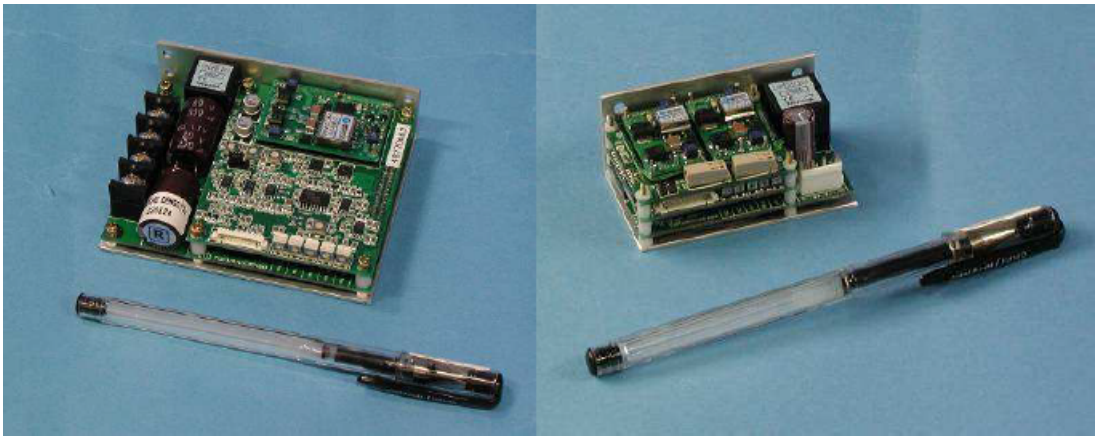


FIGURE C.10: Motor drivers: TD12770-48W10 (left) TD12770-48W05 (right).

Specifications of upper and lower body joints can be seen in Table C.3 and C.4, respectively.

TABLE C.1: Specifications of the HRP interface board.

Function	Description	Channels
A/D converter	Resolution: 12 bits Conversion time: 1.6μs (max) Input: $\pm 10V$ or $0 \sim 5v$	16
D/A converter	Resolution: 12 bits Conversion time: 30μs (typ) Output: $\pm 10V$	16
Encoder counters	Bandwidth: 4.0 MHz Resolution: 24 bits Supports TTL and Line Receiver	16
PIO ports	TTL	IN 16 / OUT 16

TABLE C.2: Specifications of the six axis force/torque sensors.

Model	IFS-67M25T50-M40BS
Dimensions	H 25 mm $\times \phi 67$ mm
Weight	250 g
Const. load	$F_x, F_y = 450N$ $F_z = 900N$ $M_x, M_y, M_z = 30Nm$
Max. load	Const. load $\times$ 400

TABLE C.3: Upper limbs joint specifications.

	Wrist			Elbow	Shoulder		
	Roll	Pitch	Yaw	Pitch	Yaw	Roll	Pitch
Motor model (Maxon®)	RE-max17 216014	RE-max17 216014	RE-max17 216014	RE30 268214	RE-max29 226806	RE-max29 226806	RE30 268214
Maximum torque [Nm]	0.40	0.40	0.40	20	2.5	7.9	20
Assigned power rating [W]	4.5	4.5	4.5	60	22	22	60
Reduction ratio	100	244	100	327	101	333	330
Movable range (Human)[deg]	-85 85	-15 55	-85 90	0 145	-80 110	-30 180	-50 180
Movable range (Robot)[deg]	-115 41	-47 47	-180 180	-10 130	-180 180	-17 196	-180 180

TABLE C.4: Lower limbs joint specifications.

	Ankle			Knee	Hip		
	Yaw	Roll	Pitch	Pitch	Pitch	Roll	Yaw
Motor model (Maxon®)	RE-max29 226806	RE35 118777	RE35 118777	RE40 148867	RE35 118777	RE35 118777	RE-max29 226806
Assigned power rating [W]	22	90	90	150	90	90	22
Reduction ratio	303	230	182	291	182	230	303
Movable range (Human)[deg]	-10	-20	-45	-130	-15	-45	-45
	20	30	20	0	125	20	45
Movable range (Robot)[deg]	-180	-45	-33	-156	-64	-22	-100
	180	25	118	0	94	22	29

TABLE C.5: Neck, trunk and waist joint specifications.

	Neck			Trunk		Waist	
	Yaw	Roll	Pitch	Pitch	Roll	Roll	Yaw
Motor model (Maxon®)	RE-max24 222053	RE-max24 222053	RE-max24 222053	RE40 263066	RE40 263066	RE35 118777	RE-max29 226806
Assigned power rating [W]	11	11	11	150	150	90	22
Reduction ratio	100	100	100	483	483	101	303
Movable range (Human)[deg]	-80	-45	-60	-30	-50	-	-40
	80	45	50	45	50	-	40
Movable range (Robot)[deg]	-180	-44	-32	-30	-50	-16	-180
	180	56	24	45	50	16	180

TABLE C.6: Segment mass information 1/2.

Segment	Weight [kg]	Segment	Weight [kg]	Segment	Weight [kg]
HEADP	0.300	LTHR	1.386	TRKB	0.231
HEADB	0.249	LTHB	0.213	HARD	6.794
HEAD	0.436	LTHI	3.988	BATTERY	4.686
RFOOT	0.586	LSHI	2.105	ELEFRONT	2.670
RAKR	1.222	LAKB	0.198	ELEBACK	5.629
RAKB	0.198	LAKR	1.222	FRONTCVR	0.817
RSHI	2.105	LFOOT	0.586	BACKCVR	1.400
RTHI	3.988	WST	2.278	ERROR	7.459
RTHB	0.213	PUB	1.845	NECKY	0.472
RTHR	1.386	TRKP	1.274	RHAND	0.350

FIGURE C.11: Six axis force/torque sensors used on the robot.

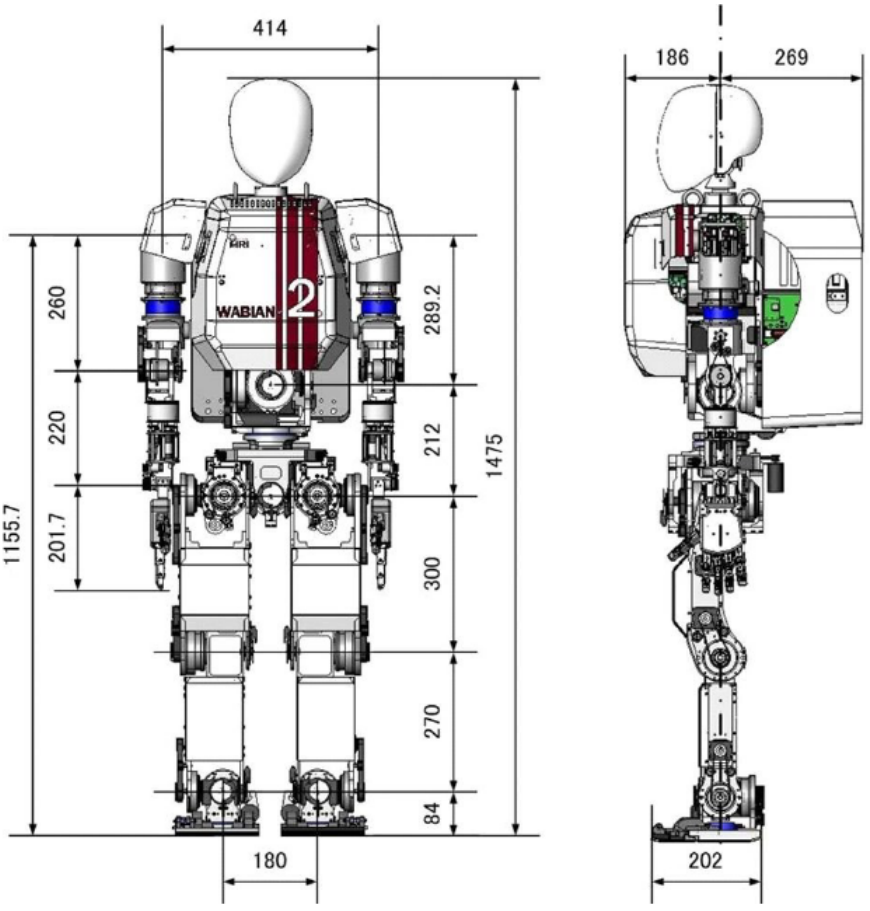


FIGURE C.12: Dimensions of WABIAN-2R.

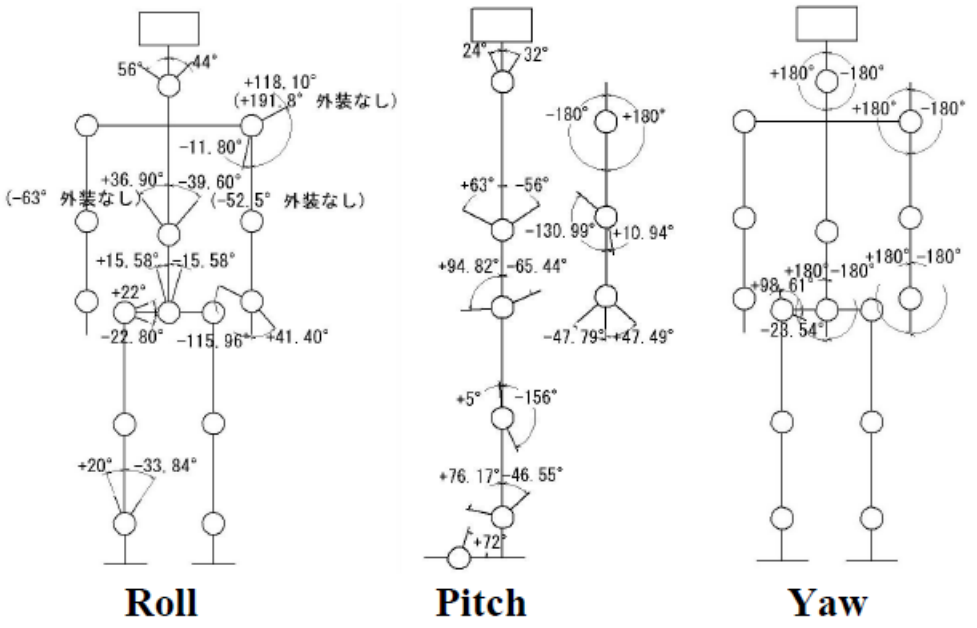


FIGURE C.13: Range of the DoF of WABIAN-2R.

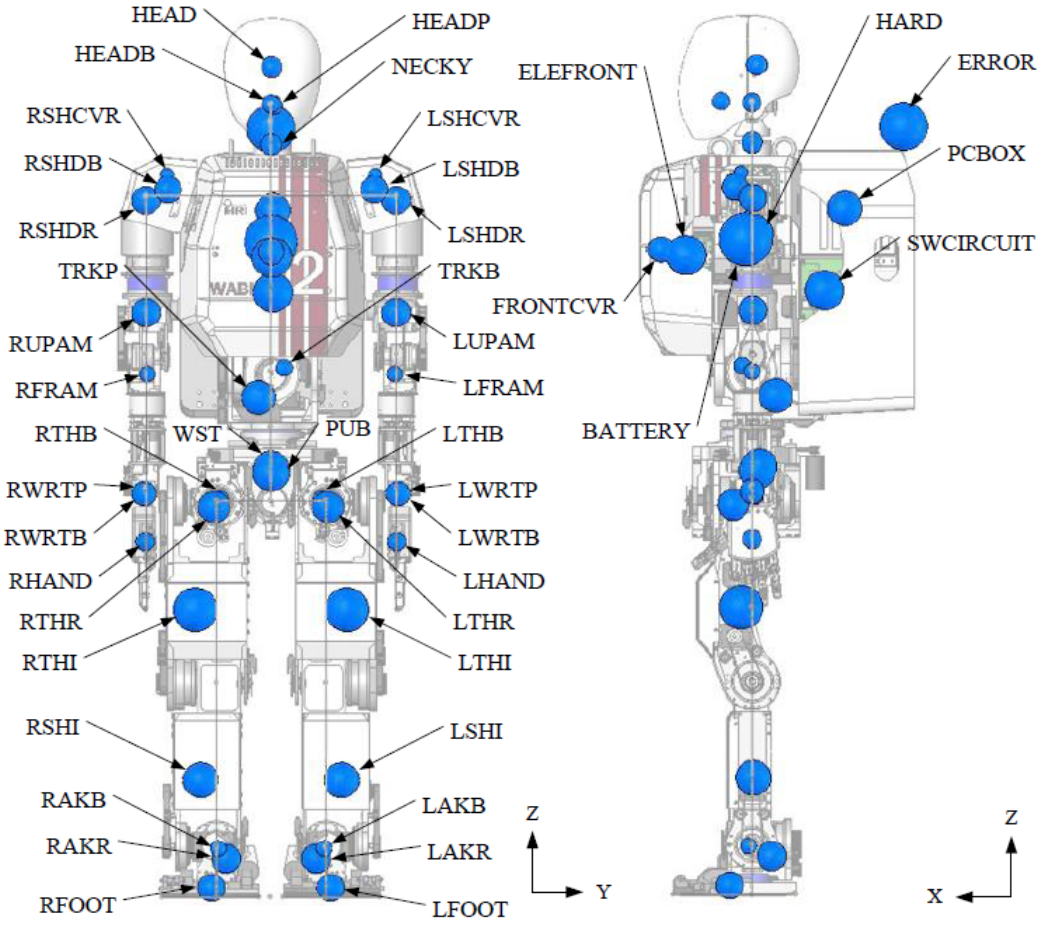


FIGURE C.14: Mass distribution of each segment of WABIAN-2R.

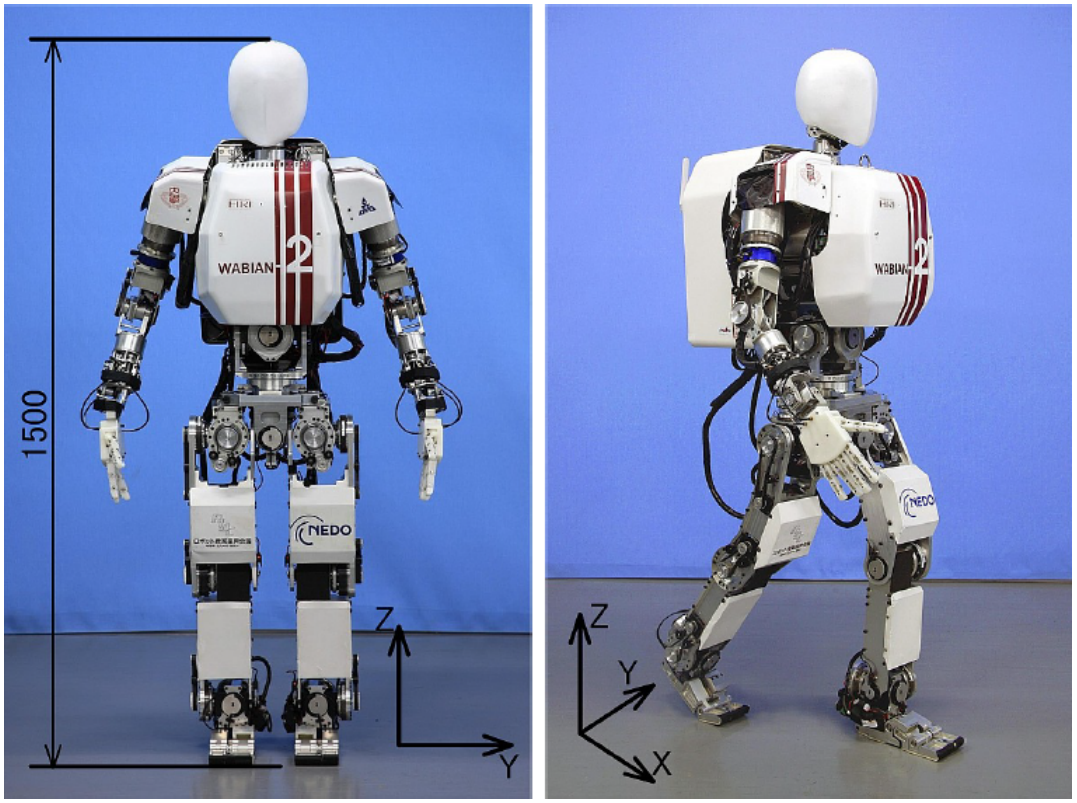


FIGURE C.15: Pictures of the actual robot, WABIAN-2R.

TABLE C.7: Segment mass information 2/2.

Segment	Weight [kg]	Segment	Weight [kg]
RWRTB	0.078	LUPAM	1.159
RWRTP	0.444	LFRAM	0.138
RFRAM	0.138	LWRTP	0.444
RUPAM	1.159	LWRTB	0.078
RSHDR	1.036	LHAND	0.350
RSHDB	0.848	RSHCVR	0.132
LSHDB	0.848	LSHCVR	0.132
LSHDR	1.036	**TOTAL**	**64.305**

Appendix D

Online Walking Gait Pattern Generator

D.1 New Linear Inverted Pendulum Based Model

To achieve a dynamically consistent heel-contact toe-off walking gait pattern, we propose a simple LIP based model to obtain an initial approximation of the feet placements and CoM trajectory of a humanoid robot. This information can be used afterwards to generate the end-effector trajectories and the reference ZMP trajectory. Then, using a multibody dynamics based gait pattern generator as in [122] or [123] and inverse kinematics, we can obtain a gait pattern for a position controlled humanoid robot, which comprises the angular references for each of its joints.

For the present research, we decided to use the LIP model [124] due to the linearity and the ease of manipulation of its equations. These are derived from the following assumptions:

- The whole system is represented by a single mass inverted pendulum, with the mass as that of the entire system, and placed at the height of the robot's CoM in the free standing position.
- The CoM motion is constrained to a horizontal plane at the height of the mass, i.e., CoM's height is constant.

- There is no torque acting between the system and the ground surface.
- There is no slip between the pendulum and the ground.

With these assumptions, the motions of the pendulum in sagittal and frontal plane can be considered as decoupled, and therefore the trajectories in each plane can be generated separately. The equation which describes these motions is:

$$\ddot{x} = \frac{g}{z}x \quad \left(\ddot{y} = \frac{g}{z}y\right) \tag{1}$$

where x and y are the position of the CoM with respect to the ground contact point of the inverted pendulum in sagittal and frontal plane respectively, g is the gravitational acceleration and z is the height of the CoM. In this model we focus on the motion on the sagittal plane, where the heel-contact toe-off motion occurs.

Integrating (1), it is possible to obtain the equations that describe the position and velocity of the CoM[125]:

$$x(t) = x_0 cosh\left(\frac{t}{k}\right) + \dot{x}_0 k \sinh\left(\frac{t}{k}\right) \tag{2}$$

$$\dot{x}(t) = \frac{x_0}{k} sinh\left(\frac{t}{k}\right) + \dot{x}_0 cosh\left(\frac{t}{k}\right) \tag{3}$$

where x_0 and $\dot{x}_0$ are the initial position and velocity of the CoM with respect to the inverted pendulum ground contact point, t is the time counted from the moment of the initial conditions and $k = \sqrt{z/g}$.

From (2) we can isolate t in order to get the time needed to reach a given final position from a given initial position and velocity. In the same way, we can isolate t from (3) to get the time needed to reach a given final velocity from a given initial position and velocity. These equations are:

$$t_{pos} = k \log\left(\frac{x_1 \pm \sqrt{k^2\dot{x}_0^2 + x_1^2 - x_0^2}}{x_0 + k\dot{x}_0}\right) \tag{4}$$

$$t_{vel} = k \log \left(\frac{k\dot{x}_1^2 \pm \sqrt{k^2(\dot{x}_1^2 - \dot{x}_0^2) + x_0^2}}{x_0 + k\dot{x}_0} \right) \tag{5}$$

where t_{pos} and t_{vel} are, the time to reach a final position x_1 or velocity $\dot{x}_1$, respectively, given initial conditions, x_0, $\dot{x}_0$.

If we equate (4) and (5), it means that we want the time to get from some initial conditions to a final position, to be the same as getting to a final velocity given the same initial conditions. Doing so, we can obtain a set of equations which relate initial to final conditions, where we can isolate any of them (position or velocity), to calculate it given that we know the other three conditions. These equations are:

$$x_0 = \pm\sqrt{k^2(\dot{x}_0^2 - \dot{x}_1^2) + x_1^2} \tag{6}$$

$$\dot{x}_0 = \pm\frac{\sqrt{k^2\dot{x}_1^2 + x_0^2 - x_1^2}}{k} \tag{7}$$

$$x_1 = \frac{x_0^2 - k^2\dot{x}_0^2 + \left(k\dot{x}_1 \pm \sqrt{x_0^2 + k^2(\dot{x}_1^2 - \dot{x}_0^2)}\right)^2}{2\left(k\dot{x}_1 \pm \sqrt{x_0^2 + k^2(\dot{x}_1^2 - \dot{x}_0^2)}\right)} \tag{8}$$

$$\dot{x}_1 = \frac{x_1^2 - x_0^2 + k^2\dot{x}_0^2 \pm x_1\sqrt{k^2\dot{x}_0^2 + x_1^2 - x_0^2}}{k\left(x_1 \pm \sqrt{k^2\dot{x}_0^2 + x_1^2 - x_0^2}\right)} \tag{9}$$

It is worth noting that, even though it is always possible to calculate a condition given the other three from (6), (7), (8) and (9), the solution is not always real, which is the case of final conditions that cannot be achieved given some initial conditions, or initial conditions that cannot produce some desired final conditions. Likewise, when calculating the motion time with (4) and (5), and given that both initial and final conditions are real, there can be different cases:

- $\{t_{pos}, t_{vel} \in \mathbb{R} : t_{pos} = t_{vel} \geq 0\}$, the motion with the given initial and final conditions is feasible, and it will take time $t = t_{pos} = t_{vel}$ for the initial conditions to reach the final conditions. (When $t_{pos} = t_{vel} = 0$, the initial conditions are the same as the final conditions.)

- $\{t_{vel} \in \mathbb{C}; t_{pos} \in \mathbb{R} : t_{pos} \geq 0\}$, the motion with the given initial conditions will reach the final position in $t = t_{pos}$, but will not reach the desired final velocity.
- $\{t_{pos} \in \mathbb{C}; t_{vel} \in \mathbb{R} : t_{vel} \geq 0\}$, the motion with the given initial conditions will reach the final velocity in $t = t_{vel}$, but will not reach the desired final position.
- $\{t_{vel}, t_{pos} \in \mathbb{C}\}$ or $\{t_{pos}, t_{vel} \in \mathbb{R} : t_{pos}, t_{vel} < 0\}$, the motion with the given initial and final conditions is not physically feasible.

Therefore, we will always seek for conditions that fulfill $\{t_{pos} = t_{vel} > 0\} \in \mathbb{R}$.

D.2 Heel-Contact Toe-Off Walking Model

We propose a model for the motion in the sagittal plane of heel-contact toe-off walking, based on the functional rockers of the foot, which are, as the name states, rocker-like motions that happen in the stance phase of a gait[126]. There are three rockers, which are named after the body part that functions as the pivot or fulcrum of the motion: heel, ankle and forefoot. The heel rocker starts with the heel contacting the ground, and finishes when the forefoot strikes the floor. Then the ankle rocker starts when the whole sole of the foot is in contact with the ground, finishing when the Center of Pressure (CoP) on the foot reaches the metatarsal heads (MTH). This leads to the start of the forefoot rocker, concluding when the foot exchange occurs, and the opposite foot's heel strikes the ground. For this model, we assumed the following:

- The CoM velocity at foot exchange is user-defined.
- There is no double support phase, i.e., the foot exchange is instantaneous.

Based on the above, we decided to model each rocker as an inverted pendulum, changing the ground contact point for each. For simplicity, the foot was considered as a rectangular board with a passive toe joint, and the different ground contact points for each phase were placed as in Fig. D.1 (b).

For the proposed model, we are using as input the velocity at the moment when the CoM is over the ankle rocker pivot, i.e., when the vector from the ground contact point

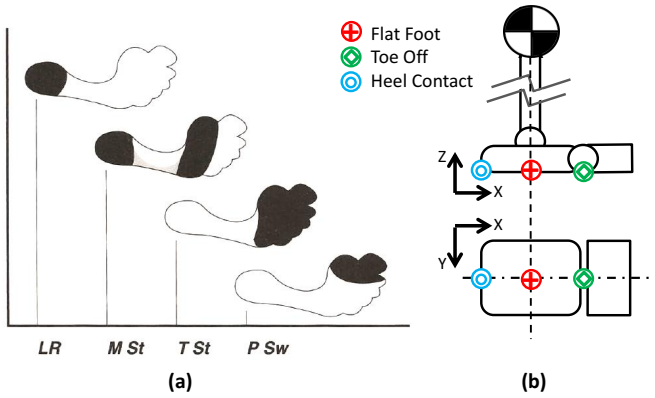


FIGURE D.1: Ground contact areas and points during the stance phase for the heel-contact toe-off motion: (a) Human foot (bottom view) [126]. (b) Model foot, side view (up) and bottom view (down).

to the CoM is completely vertical. We will call this point *apex*, as in [122]. With that in mind, we will divide and analyze the motion as follows:

- Phase I will start from the stance foot's apex and will end with the CoM over the same foot's MTH.
- Phase II will begin with the CoM over the same foot's MTH and end with the same foot's toe-off, which will coincide with the opposite foot's heel-contact (foot exchange point).
- Phase III will last from the opposite foot's heel-contact to that same foot's forefoot contacting the ground.
- Phase IV will be from the foot's forefoot contacting the ground to the new apex.

Consistent with the above phases, the time, initial position and velocity, and final position and velocity will be named with each phase's roman number: t_I, x_{0_I}, $\dot{x}_{0_I}$, x_{1_I}, $\dot{x}_{1_I}$, t_{II}, $x_{0_{II}}$, etc. (Fig. D.2)

It should be noted that regardless of the order of the present analysis, as long as the motion is physically feasible and the initial conditions and necessary velocities (apex and/or exchange point velocities) are known, it does not matter from which phase the calculations are started.

Appendix D

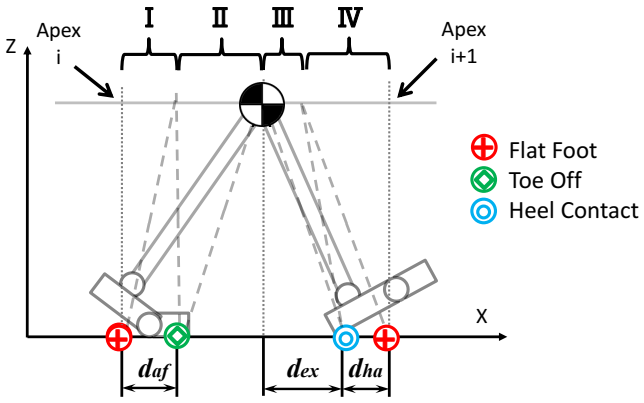


FIGURE D.2: Phases for the analysis of the motion.

Phase I: Motion from Initial Apex to CoM over Metatarsal Heads

In this phase, we know the initial velocity, $\dot{x}_{0_I} = v_{apex_i}$, as it is an input, and the initial position, $x_{0_I} = 0$, as the CoM is at an apex, and its position is measured with respect to the ground contact point. Also, we know the final position, $x_{1_I} = d_{af}$, where d_{af} is the distance between the ankle and the forefoot rockers' pivot points. This point was chosen to simplify the calculations, as it will make the initial position of the next phase to be zero. With these three conditions it is possible to calculate the motion time from (4), and the final velocity from (3):

$$t_I = k \log\left(\frac{d_{af} \pm \sqrt{k^2 v_{apex_i}^2 + {d_{af}}^2}}{k\ v_{apex_i}}\right) \tag{10}$$

$$\dot{x}_{1_I} = v_{apex_i} cosh\left(\frac{t_I}{k}\right) \tag{11}$$

It is worth noting that the final position can be placed before or after the point used here (d_{af}), as long as the effects that it will have on the motion of this and the next phase are kept in mind.

Phase II: Motion from CoM over Metatarsal Heads to Toe-off (Foot Exchange)

For this phase, we know the initial position and velocity, where the position is calculated from the final position of the previous motion, minus the distance between the ankle

and forefoot rocker pivots, $x_{0_{II}} = x_{1_I} - d_{af}$ (in the present case $x_{0_{II}} = 0$), and the initial velocity is the final velocity of the previous motion, $\dot{x}_{0_{II}} = \dot{x}_{1_I}$. For the final conditions, it is necessary to define either the final velocity $\dot{x}_{1_{II}}$ or position $x_{1_{II}}$. As we are using velocities as inputs, we decided to set the final velocity as the exchange point velocity $\dot{x}_{1_{II}} = v_{ex}$, as the foot exchange occurs in the end of this phase. With these three conditions it is possible to calculate the motion time from (5), and the final position from (2):

$$t_{II} = k \log \left(\frac{k\ v_{ex}^2 \pm \sqrt{k^2(v_{ex}^2 - \dot{x}_{1_I}^2) + (x_{1_I} - d_{af})^2}}{(x_{1_I} - d_{af}) + k\dot{x}_{1_I}} \right) \tag{12}$$

$$x_{1_{II}} = (x_{1_I} - d_{af}) cosh\left(\frac{t_{II}}{k}\right) + \dot{x}_{1_I} k \sinh\left(\frac{t_{II}}{k}\right) \tag{13}$$

To check for motion feasibility, (8) and (4) can be used, to see if $t_{pos} = t_{vel}$. In this case, we found empirically that for a feasible motion, $v_{ex} > \dot{x}_{1_I}$ is enough.

Phase III: Motion from Opposite Foot Heel-contact to Forefoot Contact

In this phase, the initial velocity is the same as the final velocity of the previous motion, $\dot{x}_{0_{III}} = v_{ex}$. As stated in the model assumptions, the motion time is fixed to a constant $t_{III} = constant$, which will be later explained in section IV.

For the initial position, we defined d_{ex} as the distance from the projection of the CoM at foot exchange to the heel contact position, so that $x_{0_{III}} = -d_{ex}$, and to define it, as the time of this motion is fixed, it is necessary to know the desired velocity of the next apex, which will occur in the next phase. Because of that we must analyze both Phase III and IV together to completely define all the conditions of both phases. So, with initial position $x_{0_{III}} = -d_{ex}$ and velocity v_{ex}, we can use (2) and (3):

$$x_{1_{III}} = -d_{ex} cosh\left(\frac{t_{III}}{k}\right) + v_{ex} k \sinh\left(\frac{t_{III}}{k}\right) \tag{14}$$

$$\dot{x}_{1_{III}} = \frac{-d_{ex}}{k} sinh\left(\frac{t_{III}}{k}\right) + v_{ex} cosh\left(\frac{t_{III}}{k}\right) \tag{15}$$

where $-d_{ex}$, $x_{1_{III}}$ and $\dot{x}_{1_{III}}$ are unknown, and therefore we need either to relate two of those variables or another function, to be able to solve a system of equations for the remaining unknowns. This relation will be obtained from the analysis of the next motion.

Phase IV: Motion from Forefoot Contact to Next Apex

Finally, for this phase we have the information of the final conditions, where the final position $x_{1_{IV}} = 0$ as it is an apex, and the final velocity $\dot{x}_{1_{IV}} = v_{apex_{i+1}}$ is an input. As for the initial conditions, we already know that the velocity corresponds to the final velocity of the previous motion, $\dot{x}_{0_{IV}} = \dot{x}_{1_{III}}$, and for the initial position we must take into account the movement of the ground contact point from the heel to the position of the ankle rocker pivot, thus $x_{0_{IV}} = x_{1_{III}} - d_{ha}$, where d_{ha} is the distance between the heel and the ankle rocker pivot. With this, we can use (7) to relate the initial velocity to the other three conditions:

$$\dot{x}_{0_{IV}} = \dot{x}_{1_{III}} = +\frac{\sqrt{k^2 v_{apex_{i+1}}^2 + (x_{1_{III}} - d_{ha})^2}}{k} \tag{16}$$

where we only take the positive value, as we are expecting a forward motion. Then we can substitute $\dot{x}_{1_{III}}$ in (15) with the right hand side of (16), and solve the system of the resulting equation and (14) to calculate $-d_{ex}$ and $x_{1_{III}}$:

$$x_{1_{III}} = d_{ha} cosh^2\left(\frac{t_{III}}{k}\right) + k v_{ex} sinh\left(\frac{t_{III}}{k}\right) \pm cosh\left(\frac{t_{III}}{k}\right) \sqrt{\left(d_{ha} sinh\left(\frac{t_{III}}{k}\right) + k v_{ex}\right)^2 - k^2 v_{apex_{i+1}}^2} \tag{17}$$

$$d_{ex} = sinh\left(\frac{t_{III}}{k}\right) \sqrt{(x_{1_{III}} - d_{ha})^2 + k^2 v_{apex_{i+1}}^2} - x_{1_{III}} cosh\left(\frac{t_{III}}{k}\right) \tag{18}$$

Here, as it can be seen from (17), there are two possible final positions for Phase III, $x_{1_{III}}$, which in turn will produce two solutions for d_{ex}. Analyzing the physical meaning of this, if $d_{ex} > 0$, it means that the heel-contact will take place in front of the position of the CoM in the foot exchange moment, which will stop the fall of the CoM, redirecting the falling force into a forward motion, which is one of the functions of the heel rocker[126]. On the other hand, if $d_{ex} \leq 0$, the heel-contact will take place behind the position of the CoM in the foot exchange moment, allowing the CoM to keep falling and having nothing to stop it, unless $d_{ex} = 0$ and $v_{ex} = 0$, in which case the motion should stop. Therefore, the positive value of d_{ex} will be chosen, and with it, the value of $x_{1_{III}}$ that produced it.

Then it only remains to assure that the generated motion is feasible, for which we should once again use (8) and (4) to see if $t_{pos} = t_{vel}$ using the initial and final conditions for Phase IV. If it is not feasible $v_{apex_{i+1}}$ should be changed.

Frontal Plane Motion

For the motion in frontal plane, we used the same method as in [122]. The steps are planned so that the CoM swings do not cross the support foot position, and they have the desired step time, obtained from the sagittal plane motion generation. The CoM motion also depends on (1), (2) and (3).

D.3 Simulations

Having the necessary equations for the different phases to obtain the foot positioning and motion of CoM for a heel-contact toe-off motion, we made kinematic simulations to test the behavior of the model. These simulations were done using the software MATLAB®. We took three scenarios into account: a constant apex velocity step, $v_{apex_i} = v_{apex_{i+1}}$, an accelerating step, $v_{apex_i} < v_{apex_{i+1}}$, and finally a decelerating step, $v_{apex_i} > v_{apex_{i+1}}$. For all the simulations, the value for t_{III} was defined inside the range 0.1-0.2s, from the assumption that the step time is around 1.0-2.0s, and the heel-contact to forefoot contact motion in humans takes about 10% of the step time. For the presented simulations, $t_{III} = 0.15s$. The tests were made under the assumption that the motion would be from apex to apex, having the velocity at those points as inputs. The velocity v_{ex} at

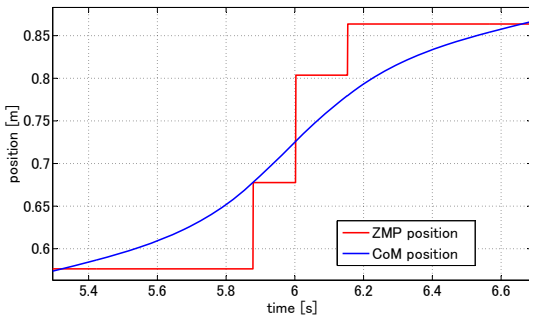


FIGURE D.3: Position of the CoM and ZMP for the constant apex velocity simulation.

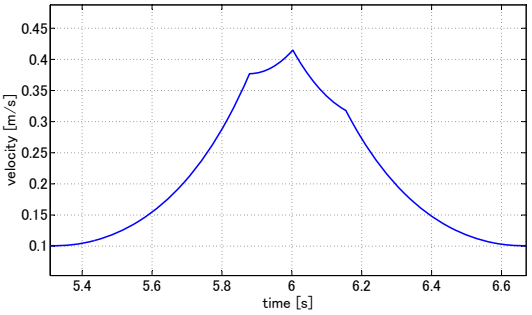


FIGURE D.4: Velocity of the CoM for the constant apex velocity simulation.

exchange point was selected so that it obeys the condition $v_{ex} > \dot{x}_{1_I}$, defining it for the simulations as $v_{ex} = 1.1 \cdot \dot{x}_{1_I}$. To prove the effectiveness of the present model, the reference ZMP which could be used to generate a walking pattern was calculated from the data from the generated motions, to see if the heel-contact toe-off motions had any effects on it:

$$ZMP_{ref} = x - \frac{z}{g}\ddot{x} \tag{20}$$

where x is the position of the CoM in the global coordinate system, and $\ddot{x}$ is the acceleration on that point.

Constant Apex Velocity

We simulated a step with constant apex velocity, i.e. $v_{apex_i} = v_{apex_{i+1}} = 0.1m/s$. In Fig. D.4, the velocity of the CoM in sagittal plane is shown, where the transition from and to each of the modeled inverted pendulums can be clearly seen. In Fig. D.3, which shows the ZMP position also in sagittal plane, the effect of the change of placement of the ground contact point of each inverted pendulum can be observed as a stair-like shape, where each "step" shows a ground contact point shift. This proves that the model is successfully coping with the effects of the change of pivot point during the heel-contact toe-off motion.

Accelerating Apex Velocity

For this test, as the calculations are not possible with a $v = 0$, we used a very small value for the initial apex, $v_{apex_i} \approx 0m/s$, and $v_{apex_{i+1}} = 0.4m/s$. Once again, the inverted

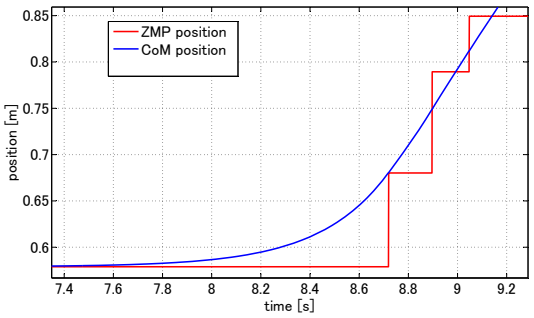


FIGURE D.5: Position of the CoM and ZMP for the accelerating apex velocity simulation.

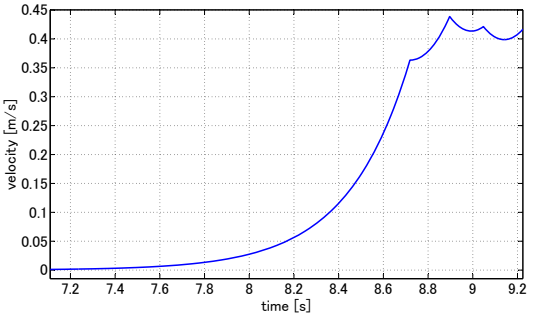


FIGURE D.6: Velocity of the CoM for the accelerating apex velocity simulation.

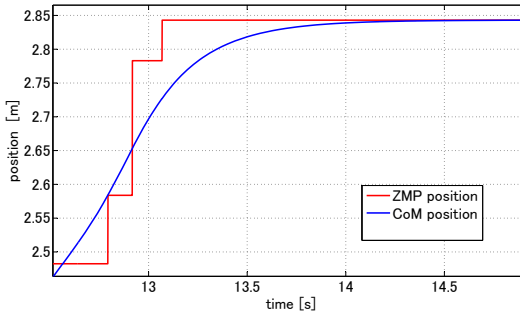


FIGURE D.7: Position of the CoM and ZMP for the decelerating apex velocity simulation.

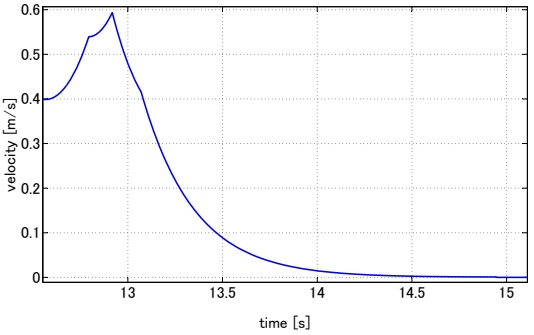


FIGURE D.8: Velocity of the CoM for the decelerating apex velocity simulation.

pendulum transitions can be clearly seen in Fig. D.6, as well as the effects of the pivot position shift to the ZMP in Fig. D.5. It is also notable that the model does not have problems to get from the desired initial to final apex velocities. A change in the step time and length with respect to the other two simulations is observed, which shows how the model tunes these parameters by itself to achieve the desired velocities.

Decelerating Apex Velocity

For this test, we used $v_{apex_i} = 0.4m/s$ and $v_{apex_{i+1}} \approx 0m/s$. The results are very similar the other simulations, and it is shown that there is no problem to generate a decelerating motion as well, reaching the desired final apex velocity and accounting for the effects of the change of pivot point during each of the pendulums.

As an output of this first stage, the positioning and timing of each foot is obtained, as well as the trajectory of the CoM. This can be seen in Fig. D.9.

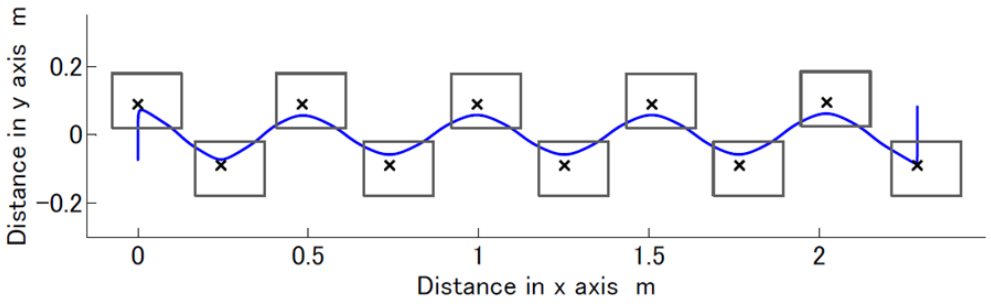


FIGURE D.9: Output of the first stage.

D.4 Multibody Dynamics Model

In this stage, just like in 3, the reference trajectories for the end-effectors (feet, hands and CoM) are defined using the parameters calculated in the previous stage, to next define the whole walking motion for the whole robot, and finally refine the references using the multibody dynamics model. In the following sections, solutions to enable heel-contact toe-off motions will be explained.

Here, the focus was on the feet trajectories, as the heel-contact toe-off motions does not have big effects on the hands trajectories.

D.4.1 Feet Reference Trajectories

For the feet trajectories, first it was now necessary to obtain not only the information about feet transitions, timings and durations, but also the timings and durations of the heel-contact to flatfoot phase and the heel-off to toe-off phase. As the model from the first stage already accounts for the heel-contact toe-off effects, this information can still be extracted from the results of the LIP model stage. Having the feet transitions, timings and durations for all the phases, it is then required to generate the trajectories for the different phases, namely heel-off to toe-off, swing and heel-contact to flatfoot phases. Each of these have different characteristics, and these were based on the physical model of the feet mechanisms of the robot (Fig. D.10), and the values can be seen in Table D.1.

About the lack of double support time, an artificial double support phase is also introduced by delaying the toe-off and accelerating the heel-contact of the feet. Also, the

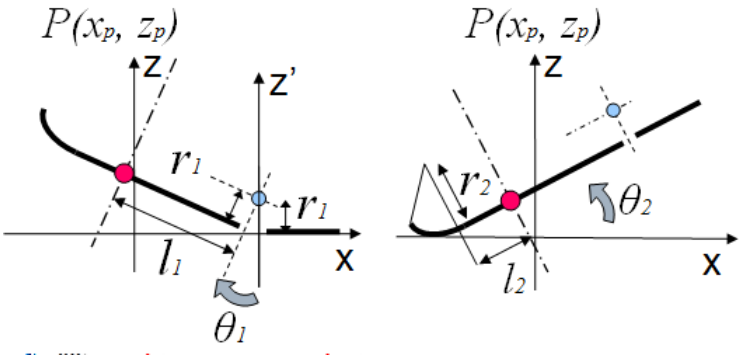


FIGURE D.10: Models based on the robot's feet for toe-off (left) and heel-contact (right).

TABLE D.1: Feet parameters

Parameter	Value [mm]
Forefoot rocker radius r_1	14.0
Ankle projection to forefoot joint l_1	82.5
Heel rocker radius r_2	30.0
Heel to ankle projection l_2	55.6

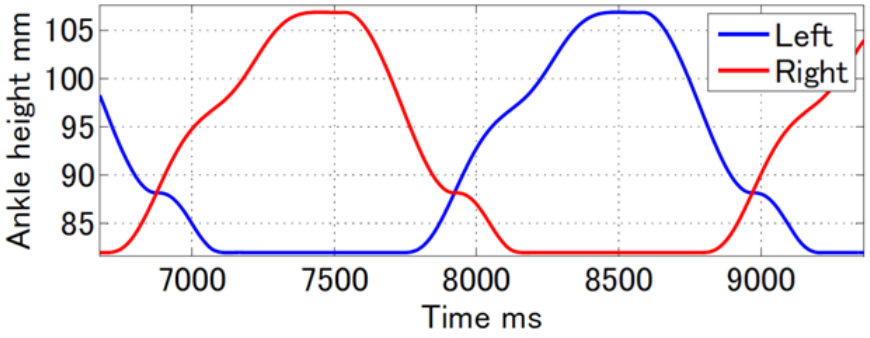


FIGURE D.11: Generated trajectory of the ankle in the z axis.

timing of the beginning of the heel-off toe-off) motion can be changed. The generated trajectory of the ankle can be seen in Fig. D.13.

D.4.2 ZMP Reference and Preview Controller

For the stabilization, the method remains the same as in chapter 3. The ZMP is calculated by:

$$ZMP_x^{ref} =_w x - \frac{z_c}{g}\ddot{x} \tag{3.15}$$

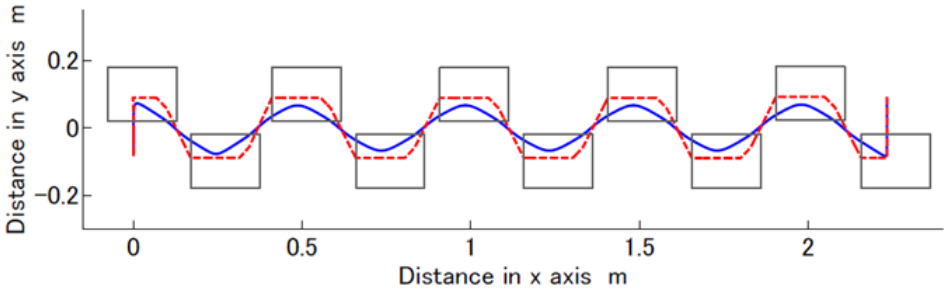


FIGURE D.12: Obtained ZMP with the new model. Note the effects of each inverted pendulum.

Then the ZMP is calculated, a low pass filter is applied to smoothen the reference. The obtained ZMP for a generated pattern can be observed in Fig. D.12, where the effect of each of the modeled inverted pendulums is clearly visible.

Finally, having the end-effector trajectories and the ZMP reference, the whole body motions are refined through the Preview Controller.

In the present thesis, a model to generate heel-contact toe-off motions based on the LIP model and the foot's functional rockers was presented. It uses apex velocities as inputs, and gives as outputs foot placements, CoM position, velocity and acceleration during the motion, and step timings and phases. For this, the stance phase of a gait was divided into four phases: from the apex to the CoM over the MTH, from CoM over the MTH to foot exchange time, from the opposite foot's heel-contact to forefoot contact, and from the forefoot contact to the next apex. The methods to calculate initial and final conditions for each phase and to verify the feasibility of the motions were presented, as well as the results of simulations using this model.

In the simulation results, it was possible to see each phase clearly in the velocity graphs, as well as the behavior of the model in each phase, proving that the model successfully emulated each of the functional rockers of the foot. Also in the graphs of the ZMP calculated from the motion data, the effect from each transition between rockers could be clearly seen, regardless of the scenario of the motion, showing that the changes of pivot points during the heel-contact toe-off motion are successfully modeled.

The generated patterns were successfully implemented on the humanoid biped robot WABIAN-2R, achieving stable heel-contact toe-off walking motions, and the possibility to change the speed of each step online.

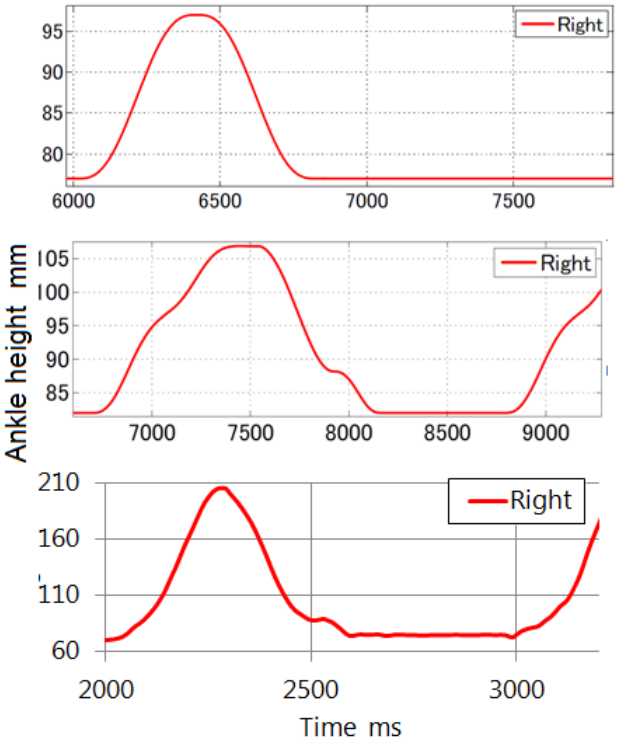


FIGURE D.13: Ankle height of previous model (up), new model (center) and human (bottom).

About the current limitations, as one of the goals is to generate more humanlike motions, it is necessary to change some assumptions. The inclusion of a double support phase and the possibility to make the time from heel-contact to forefoot contact variable are being considered. Also, the exchange point velocity is currently being defined given some empirical conditions, which could possibly be based on gait parameters, allowing to have other desired behaviors, or could be eliminated to make the method more robust.

Another important required modification is to relax the COM height constraint, in order to make the COM height variable and realize more humanlike motions with completely stretched knees. Also, the humanlikeness feet trajectories for the multibody dynamics model could be improved (Fig. D.13).

早稲田大学　博士（工学）　学位申請　研究業績書

(List of research achievements for application of doctorate (Dr. of Engineering), Waseda University)

氏 名(Full Name)　MINAMI SHIGUEMATSU, Yukitoshi　　印(seal or signature　　)

（As of July, 2019）

種 類 別 (By Type)	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）(theme, journal name, date & year of publication, name of authors inc. yourself)			
1. 論文				
○	Effects of Walking Style and Symmetry on the Performance of Localization Algorithms for a Biped Humanoid Robot	2019 IEEE/SICE International Symposium on System Integration	Jan. 2019 (published)	**Y. Minami Shiguematsu**, M. Brandao, and A. Takanishi
○	Effects of Biped Humanoid Robot Walking Gaits on Sparse Visual Odometry Algorithms	2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)	Nov. 2018 (published)	**Y. Minami Shiguematsu**, M. Brandao, K. Hashimoto and A. Takanishi
	Development of a low-cost smart home system using wireless environmental monitoring sensors for functionally independent elderly people	2017 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 153-158	Dec. 2017 (published)	D. Zhang, W. Kong, R. Kasai, Z. Gu, **Y. Minami Shiguematsu**, S. Cosentino, S. Sessa and A. Takanishi
	Material Recognition CNNs and Hierarchical Planning for Biped Robot Locomotion on Slippery Terrain	2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids), pp. 81-88	Nov. 2016 (published)	M. Brandao, **Y. Minami Shiguematsu**, K. Hashimoto and A. Takanishi
○	Heel-Contact Toe-Off Walking Pattern Generator Based on the Linear Inverted Pendulum	International Journal of Humanoid Robotics, vol. 13, no. 01, p. 1650002	Mar. 2016 (published)	**Y. M. Shiguematsu**, P. Kryczka, K. Hashimoto, H.-O. Lim, and A. Takanishi
○	Heel-contact Toe-off Walking Model Based on the Linear Inverted Pendulum	5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob), pp. 221-226	Aug. 2014 (published)	**Y. Minami**, P. Kryczka, K. Hashimoto H.-O. Lim, and A. Takanish
	Towards Dynamically Consistent Real-time Gait Pattern Generation for Full-size Humanoid Robots	2013 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 1408-1413	Dec. 2014 (published)	P. Kryczka, **Y. Minami**, P. Kormushev, K. Hashimoto, H.-O. Lim, and A. Takanishi

早稲田大学　博士（工学）　学位申請　研究業績書

(List of research achievements for application of doctorate (Dr. of Engineering), Waseda University)

種 類 別 By Type	題名、　発表・発行掲載誌名、　発表・発行年月、　連名者（申請者含む）(theme, journal name, date & year of publication, name of authors inc. yourself)			
2. 発表 ○	Towards a Sensorimotor System Based on that of Humans to Study its Effects on Walking Stabilization	2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2015)	Oct. 2015 (poster)	**Y. Minami Shiguematsu**, A. Morell, K. Hashimoto, J. Toledo, L. Acosta, Atsuo Takanishi
	人体運動シミュレータとしての２足ヒューマノイドロボットの開発（第 17 報 : 歩行中の視線安定を維持する頭部姿勢安定化モデル）	第 31 回 日本ロボット学会 学術講演会	Sep. 2013 (oral presentation)	P. Kryczka, **Y. Minami**, T. Otani, K. Hashimoto, E. Falotico, C. Laschi, P. Dario, A. Berthoz, H.-O. Lim, and A. Takanishi